



**TUGAS AKHIR – SS141501**

**PEMODELAN KASUS PENYAKIT DIARE PADA  
BALITA DI KOTA SURABAYA MENGGUNAKAN  
REGRESI NONPARAMETRIK SPLINE TRUNCATED**

**MUHAMMAD SYAUQI KHUDZAIFI  
NRP 1313 100 055**

**Dosen Pembimbing  
Dra. Madu Ratna, M.Si  
Prof. Dr. Drs. I Nyoman Budiantara, M.Si**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2017**



**TUGAS AKHIR – SS141501**

**PEMODELAN KASUS PENYAKIT DIARE PADA  
BALITA DI KOTA SURABAYA MENGGUNAKAN  
REGRESI NONPARAMETRIK SPLINE TRUNCATED**

**MUHAMMAD SYAUQI KHUDZAIFI  
NRP 1313 100 055**

**Dosen Pembimbing  
Dra. Madu Ratna, M.Si  
Prof. Dr. Drs. I Nyoman Budiantara, M.Si**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2017**



**FINAL PROJECT – SS141501**

**MODELING OF TODDLER’S DIARRHEA DISEASE  
CASES IN SURABAYA CITY USING TRUNCATED  
SPLINE NONPARAMETRIC REGRESSION**

**MUHAMMAD SYAUQI KHUDZAIFI  
NRP 1313 100 055**

**Supervisor  
Dra. Madu Ratna, M.Si  
Prof. Dr. I Nyoman Budiantara, M.Si**

**UNDERGRADUATE PROGRAM  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS AND NATURAL SCIENCES  
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY  
SURABAYA 2017**

## LEMBAR PENGESAHAN

### PEMODELAN KASUS PENYAKIT DIARE PADA BALITA DI KOTA SURABAYA MENGUNAKAN REGRESI NONPARAMETRIK SPLINE TRUNCATED

#### TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada  
Program Studi Sarjana Departemen Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember

Oleh :

**Muhammad Syauqi Khudzaifi**  
NRP. 1313 100 055

Disetujui oleh Pembimbing:  
Dra. Madu Ratna, M.Si.  
NIP. 19590109 198603 2 001

Prof. Dr. Drs. I Nyoman Budiantara, M.Si.  
NIP. 19650603 198903 1 003



Diketahui,  
Kepala Departemen  
**Dr. Suhartono**  
NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

# **PEMODELAN KASUS PENYAKIT DIARE PADA BALITA DI KOTA SURABAYA MENGGUNAKAN REGRESI NONPARAMETRIK SPLINE TRUNCATED**

**Nama Mahasiswa : Muhammad Syauqi Khudzaifi**  
**NRP : 13 13 100 055**  
**Departemen : Statistika**  
**Dosen Pembimbing 1 : Dra. Madu Ratna, M.Si**  
**Dosen Pembimbing 2 : Prof. Dr. Drs. I Nyoman Budiantara, M.Si**

## **Abstrak**

*Penyakit diare saat ini masih merupakan salah satu penyebab utama morbiditas dan mortalitas pada anak di seluruh dunia terutama di negara-negara berkembang. Berdasarkan data dari UNICEF 2015 jumlah kematian balita yang disebabkan diare di dunia mencapai 525.977 jiwa, angka ini menunjukkan hampir mencapai 10% dari total kematian balita di dunia. Surabaya sebagai salah satu kota besar di Indonesia ternyata tingkat kesakitan diare yang ditemukan masih tinggi. Berdasarkan data Dinas Kesehatan Kota Surabaya, penyakit diare yang ditangani di Kota Surabaya pada tahun 2015 sebanyak 65.447 kasus. Terdapat 4 faktor yang diduga mempengaruhi jumlah kasus diare pada balita, yaitu kepadatan penduduk, presentase rumah tangga berperilaku hidup bersih sehat (PHBS), presentase penduduk dengan akses jamban sehat, presentase bayi dengan ASI eksklusif 0-6 bulan. Pemodelan jumlah kasus diare pada balita menggunakan regresi nonparametrik spline truncated karena hubungan antara presentase kasus diare balita di Kota Surabaya dengan faktor yang diduga mempengaruhinya tidak membentuk pola tertentu. Berdasarkan hasil analisis, keempat faktor yang digunakan berpengaruh signifikan terhadap presentase kasus diare balita di Kota Surabaya. Model regresi nonparametrik spline truncated menghasilkan koefisien determinasi sebesar 91,04%.*

**Kata Kunci : Kota Surabaya, Penyakit Diare Balita, Regresi Nonparametrik Spline Truncated**

*(Halaman ini sengaja dikosongkan)*

# **MODELING OF TODDLER'S DIARRHEA DISEASE CASES IN SURABAYA CITY USING TRUNCATED SPLINE NONPARAMETRIC REGRESSION**

**Name** : Muhammad Syauqi Khudzaifi  
**Student Number** : 13 13 100 055  
**Department** : Statistics  
**Supervisor 1** : Dra. Madu Ratna, M.Si  
**Supervisor 2** : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

## **Abstract**

*Diarrhea still becoming one of the main reason of morbidity and mortality worldwide, particularly in developing countries. According to UNICEF in 2015, the total mortality among toddlers which caused by diarrhea reaches number 525.977 lives. This number contributes to 10% of the total mortality among toddlers worldwide. Surabaya as one of the biggest city in Indonesia surprisingly still has a high morbidity of diarrhea disease. According to Health and Public Care Department in Surabaya, there are 65.477 cases of diarrhea in 2015. There are 4 factors that allegedly affect the high number of diarrhea cases in Surabaya, which are population density, percentage of household which already implementing clean and healthy life, percentage of population with good and healthy latrines system, and percentage of baby with exclusive infant nutrition in 0 until 6 months. In this final project, case modeling of diarrhea disease among toddler will be solved by using nonparametric regression spline truncated. This may caused by no certain pattern that can be formed among these variables used. Based on the analysis, all factors has significantly impacted to total percentage of diarrhea case among toddlers in Surabaya. The regression model results 91,04% of correlation.*

**Keywords:** *Diarrhea disease in toddler, Nonparametric Truncated Spline Regression, Surabaya City*

*(Halaman ini sengaja dikosongkan)*



## KATA PENGANTAR

Puji syukur Alhamdulillah senantiasa penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “ **PEMODELAN KASUS PENYAKIT DIARE PADA BALITA DI KOTA SURABAYA MENGGUNAKAN REGRESI NONPARAMETRIK SPLINE TRUNCATED** ” dengan lancar dan tepat waktu.

Keberhasilan penyusunan Tugas Akhir ini tidak lepas dari partisipasi berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dra. Madu Ratna, M.Si, Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si, Bapak Dr. Ir. Setiawan M.Si dan Ibu Erma Oktania Permatasari, S.Si, M.Si selaku dosen pembimbing dan dosen penguji atas semua bimbingan, waktu, semangat dan perhatian yang telah diberikan sehingga Tugas Akhir ini dapat diselesaikan dengan baik.
2. Bapak Dr. Suhartono selaku Kepala Departemen Statistika ITS yang telah memberikan fasilitas dalam kelancaran Tugas Akhir ini
3. Kedua orangtua dan seluruh saudara-saudara, atas segala doa dan perhatian kepada penulis.
4. Mahasiswa Jurusan Statistika Angkatan 2013 atas semangat yang diberikan pada penulis.
5. Semua pihak yang telah membantu penulis yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa Tugas Akhir ini masih terdapat kekurangan, oleh karena itu kritik dan saran sangat diharapkan. Semoga Tugas Akhir ini dapat memberikan manfaat baik bagi penulis, pembaca, dan semua pihak.

Surabaya, Juli 2017

Penulis

*(Halaman ini sengaja dikosongkan)*

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	v
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	ix
<b>KATA PENGANTAR</b> .....	xi
<b>DAFTAR ISI</b> .....	xiii
<b>DAFTAR GAMBAR</b> .....	xviii
<b>DAFTAR TABEL</b> .....	xix
<b>DAFTAR LAMPIRAN</b> .....	xxi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian .....	5
1.4 Manfaat Penelitian .....	5
1.5 Batasan Masalah .....	6
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Statistika Deskriptif.....	7
2.2 Analisis Regresi .....	7
2.2.1 Regresi Nonparametrik .....	8
2.2.2 Regresi Nonparametrik Spline Truncated .....	8
2.3 Estimasi Parameter.....	9
2.4 Pemilihan Titik Knot Optimal.....	10
2.5 Pengujian Parameter Model Regresi .....	11
2.5.1 Pengujian Secara Serentak.....	12
2.5.2 Pengujian Secara Parsial/Individu .....	13
2.6 Koefisien Determinasi.....	13
2.7 Pengujian Asumsi Residual Model Regresi .....	14
2.7.1 Asumsi Identik .....	14
2.7.2 Asumsi Independen.....	15
2.7.3 Asumsi Distribusi Normal .....	16
2.8 Skenario Model .....	16
2.9 Diare.....	17

### **BAB III METODOLOGI PENELITIAN**

3.1 Sumber Data.....	19
3.2 Variabel Penelitian.....	19
3.3 Struktur Data.....	21
3.4 Langkah Analisis.....	21
3.5 Diagram Alir .....	23

### **BAB IV ANALISIS DAN PEMBAHASAN**

4.1 Gambaran Umum Jumlah Penderita Diare Balita di Kota Surabaya dan Faktor yang Diduga Mempengaruhinya .....	25
4.2 <i>Scatterplot</i> Jumlah Penderita Diare Balita dengan Faktor yang Diduga Mempengaruhinya....	29
4.3 Pemodelan Jumlah Penderita Diare Balita Menggunakan Regresi Nonparametrik Spline Truncated .....	32
4.4 Pemilihan Titik Knot Optimum.....	32
4.4.1 Regresi Nonparametrik Spline Truncated dengan Satu Knot.....	32
4.4.2 Regresi Nonparametrik Spline Truncated dengan Dua Knot.....	33
4.4.3 Regresi Nonparametrik Spline Truncated dengan Tiga Knot .....	35
4.4.4 Regresi Nonparametrik Spline Truncated dengan Kombinasi Knot .....	37
4.5 Pemilihan Model Terbaik.....	39
4.6 Penaksir Parameter Model Kasus Diare Balita di Kota Surabaya.....	39
4.7 Signifikansi Parameter.....	40
4.7.1 Uji Serentak.....	40
4.7.2 Uji Individu .....	41
4.8 Pengujian Asumsi Residual .....	42
4.8.1 Asumsi Identik .....	42
4.8.2 Asumsi Independen .....	42
4.8.3 Asumsi Distribusi Normal.....	43

4.8 Intepretasi Model .....	44
<b>BAB V KESIMPULAN</b>	
5.1 Kesimpulan .....	51
5.2 Saran .....	52
<b>DAFTAR PUSTAKA</b> .....	53
<b>LAMPIRAN</b> .....	55
<b>BIODATA PENULIS</b> .....	83

*(Halaman ini sengaja dikosongkan)*

## DAFTAR GAMBAR

	Halaman
<b>Gambar 3.1</b> Diagram Alir Penelitian.....	23
<b>Gambar 4.1</b> Presentase Kasus Diare Balita tiap Kecamatan di Kota Surabaya 2016.....	28
<b>Gambar 4.2</b> <i>Scatterplot</i> Antara Jumlah Penderita Diare Balita dengan Kepadatan Penduduk.....	29
<b>Gambar 4.3</b> <i>Scatterplot</i> Antara Jumlah Penderita Diare Balita dengan Rumah Tangga ber-PHBS.....	30
<b>Gambar 4.4</b> <i>Scatterplot</i> Antara Jumlah Penderita Diare Balita dengan Penduduk dengan Akses Jamban Sehat ..	30
<b>Gambar 4.5</b> <i>Scatterplot</i> Antara Jumlah Penderita Diare Balita dengan Bayi Mendapatkan ASI Eksklusif 0-6 bulan .....	31
<b>Gambar 4.6</b> <i>Scatterplot</i> antara residual dengan <i>observation order</i> .....	43
<b>Gambar 4.7</b> Plot Residual.....	44

*(Halaman ini sengaja dikosongkan)*



## DAFTAR TABEL

	Halaman
<b>Tabel 2.1</b> Analisis ragam (ANOVA) Uji Parameter.....	13
<b>Tabel 3.1</b> Variabel Penelitian .....	19
<b>Tabel 3.2</b> Struktur Data .....	21
<b>Tabel 4.1</b> Gambaran Umum Kasus Diare Balita di Surabaya dan Faktor-Faktor yang Diduga Memengaruhinya .....	26
<b>Tabel 4.2</b> Pemilihan Titik Knot Optimum dengan Satu Titik Knot .....	33
<b>Tabel 4.3</b> Pemilihan Titik Knot Optimum dengan Dua Titik Knot .....	34
<b>Tabel 4.4</b> Pemilihan Titik Knot Optimum dengan Tiga Titik Knot .....	36
<b>Tabel 4.5</b> Pemilihan Titik Knot Optimum dengan Kombinasi Titik Knot .....	38
<b>Tabel 4.6</b> Nilai GCV Masing-Masing Knot .....	39
<b>Tabel 4.7</b> ANOVA Model Regresi Spline Truncated.....	40
<b>Tabel 4.8</b> Hasil Pengujian Parameter Secara Individu .....	41
<b>Tabel 4.9</b> Hasil Pengujian <i>Glejser</i> .....	42
<b>Tabel 4.10</b> Prediksi Model Optimum, <i>Middle</i> , Pesimis.....	50

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1.</b> Data Penelitian .....	55
<b>Lampiran 2.</b> Program GCV 1 Knot.....	56
<b>Lampiran 3.</b> Program GCV 2 Knot.....	58
<b>Lampiran 4.</b> Program GCV 3 Knot.....	61
<b>Lampiran 5.</b> Program GCV Kombinasi Knot .....	64
<b>Lampiran 6.</b> Program Penaksiran dan Signifikansi Parameter.....	72
<b>Lampiran 7.</b> Program Uji Glejser.....	75
<b>Lampiran 8.</b> Output Penaksiran dan Signifikansi Parameter.....	77
<b>Lampiran 9.</b> Output Uji Glejser .....	79
<b>Lampiran 10.</b> Surat Pernyataan Data .....	81

*(Halaman ini sengaja dikosongkan)*

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Kelangsungan hidup berbangsa dan bernegara bergantung pada kualitas anak-anak pada masa kini sebagai generasi penerus bangsa dan negara. Menjadi sebuah negara yang kuat dan sejahtera adalah dambaan setiap bangsa di dunia. Hal ini dapat terwujud apabila suatu bangsa mempunyai kualitas yang baik diberbagai bidang kehidupan seperti pada sektor ekonomi, pendidikan, maupun kesehatan sehingga dapat dijadikan modal dalam persaingan dengan negara lain. Namun pada kenyataannya khususnya pada kualitas sektor kesehatan masih menjadi salah satu hal yang perlu diwaspadai dan harus terus ditingkatkan kualitasnya karena berbagai penyakit dapat menular dan persebarannya pun sangat cepat.

Anak-anak khususnya yang masih balita masih rentan terjangkit berbagai penyakit baik yang disebabkan oleh virus, bakteri, maupun faktor lingkungan. Salah satu penyakit yang rentan menyerang balita adalah penyakit diare. Diare mengacu pada kehilangan cairan dan elektrolit secara berlebihan yang terjadi dengan bagian feces tidak terbentuk (Nettina, 2001). Diare adalah kondisi frekuensi defekasi yang lebih dari 3 kali sehari, serta konsistensi feses yang cair (Widjaja, 2002). Penyakit diare masih menjadi salah satu masalah kesehatan masyarakat yang penting, karena angka kesakitan dan kematian anak di berbagai negara belahan dunia termasuk Indonesia masih tinggi. Berdasarkan data dari United Nation Children's Fund (UNICEF) 2015, jumlah kematian balita yang disebabkan diare di dunia mencapai 525.977 jiwa. Angka ini menunjukkan masih tingginya tingkat mortalitas pada balita yang disebabkan oleh penyakit diare, karena hampir mencapai 10% dari total kematian balita di dunia. Indonesia merupakan salah satu negara berkembang dengan angka kejadian penyakit diare yang tinggi karena tingginya morbiditas dan mortalitas (Magdarina, 2010). Jumlah

penyakit diare ditangani di Indonesia juga masih tinggi, berdasarkan data pada profil kesehatan Indonesia tahun 2015 diketahui jumlahnya adalah sebesar 4.017.861 jiwa.

Program Sustainable Development Goals (SDG's) yang terbentuk tahun 2015 terdiri dari 17 tujuan yang akan dicapai selama 15 tahun sejak program ini ditetapkan. Salah satunya adalah target di tahun 2030 mengurangi jumlah kematian balita serendah-rendahnya menjadi sebesar 25 per 1.000 kelahiran hidup. Selain itu terdapat target lain yang berhubungan dengan sanitasi yaitu terjangkau akses air minum untuk semua orang, akses ke sanitasi yang memadai, mengakhiri buang air besar terbuka, memberikan perhatian khusus pada kebutuhan perempuan dan anak serta meningkatkan kualitas air dengan mengurangi pencemaran. Tahun 2015 angka kematian balita di Indonesia 27 kasus kematian per 1000 kelahiran hidup, artinya negara Indonesia cukup berhasil. Namun keberhasilan ini harus tetap diwaspadai karena diare sampai sekarang masih menjadi masalah kesehatan masyarakat dan sering timbul dalam bentuk Kejadian Luar Biasa (KLB) disertai angka kematian yang tinggi sehingga harus selalu diantisipasi agar target SDG's ini dapat terlaksana dengan baik. Berdasarkan Kajian Morbiditas Diare 2012 yang dilakukan oleh Departemen Kesehatan RI menunjukkan angka kesakitan diare pada semua umur sebesar 214 per 1.000 penduduk dan pada balita 900 per 1.000 penduduk Kematian diare pada balita 75,3 per 100.000 balita dan semua umur 23,2 per 100.000 penduduk.

Surabaya merupakan salah satu kota besar di Indonesia yang memiliki tata ruang kota bagus dan memperoleh berbagai penghargaan. Salah satu penghargaan yang diperoleh oleh Kota Surabaya adalah Adipura Kencana sebagai kota metropolitan paling bersih di Indonesia tiga tahun berturut-turut sejak tahun 2012 merujuk pada kota yang berhasil melampaui batas pencapaian pengendalian pencemaran air, udara, pengelolaan tanah, perubahan iklim, sosial, ekonomi, dan keanekaragaman hayati. Namun hal ini ternyata belum mencerminkan kualitas

kesehatan yang sudah baik bagi Kota Surabaya, karena banyak dijumpai kasus terjadinya penyakit diare khususnya pada balita.

Berdasarkan data dari Dinas Kesehatan Kota Surabaya, penyakit diare yang ditangani di Kota Surabaya pada tahun 2015 sebanyak 65.447 kasus dari 60.960 perkiraan kasus yang ada atau sebesar 107,36% dan jika berdasarkan kelompok usia khususnya balita, penyakit diare yang ditangani adalah sebanyak 16.234 jiwa yang tersebar diberbagai kecamatan di Kota Surabaya. Hal ini menandakan bahwa kasus kesakitan penyakit diare pada balita di Surabaya masih sangat tinggi. Jika dilihat dari jumlah kasus pada tahun-tahun sebelumnya jumlah penderita diare cenderung tidak jauh berbeda sehingga dapat dikatakan penyakit diare di Surabaya merupakan penyakit endemik. Berbagai upaya penanganan, seperti penyuluhan tentang kebersihan lingkungan, penyuluhan tentang pemilahan sampah dan lain-lain selalu dilakukan terjadwal pada posyandu serta program kerja bakti yang diselenggarakan dari dinas kesehatan, namun upaya-upaya tersebut masih belum memberikan hasil yang memuaskan. Angka kematian yang tinggi akibat diare akan berdampak negatif pada kualitas pelayanan kesehatan karena angka kematian anak (AKA) merupakan salah satu indikator untuk menilai derajat kesehatan yang optimal. Kurang berhasilnya usaha dalam proses pencegahan diare merupakan salah satu faktor yang harus diperhatikan karena jika upaya pencegahan tidak ditanggulangi dengan baik, maka peningkatan penyakit diare akan semakin meningkat (Depkes, 2010).

Penelitian sebelumnya mengenai kasus diare di Kota Surabaya telah dilakukan oleh Ernawati(2015) menggunakan metode *Geographically Weighted Negative Binomial Regression*. Penelitian tersebut menggunakan data tahun 2014 dengan pemodelan dan pemetaan terhadap kasus penyakit diare yang ditinjau dari letak lokasi kecamatan di Kota Surabaya. Hasil dari penelitian ini menunjukkan bahwa variabel kepadatan penduduk, rasio sarana kesehatan puskesmas, dan rasio dokter umum berpengaruh signifikan terhadap kasus diare di seluruh kecamatan

di Kota Surabaya. Untuk variabel rumah tangga dengan jamban sehat dan tempat sampah sehat, dan rumah tangga berperilaku hidup bersih dan sehat (PHBS) hanya berpengaruh signifikan pada beberapa kecamatan di Kota Surabaya terhadap jumlah kasus diare. Kelemahan dari penelitian ini adalah variabel respon yang digunakan mencakup semua kelompok umur padahal jumlah penderita terbanyak dan rentan adalah hanya pada kelompok umur balita, selain itu penelitian ini hanya untuk mengetahui signifikansi dari variabel prediktor. Model yang terbentuk juga tidak memperhatikan aspek nonlinear dari data. Penelitian lain dilakukan oleh Ayuningrum(2015) mengenai analisis faktor sanitasi dan sumber minum pada insiden diare di Jawa Timur dengan regresi logistik biner. Kelemahan dari penelitian ini adalah hanya ingin mengetahui faktor signifikansi penyebab diare dari segi sanitasi dan sumber minum saja tanpa dilakukan pemodelan. Sedangkan untuk penelitian sebelumnya yang menggunakan metode yang sama yaitu menggunakan metode regresi nonparametrik Spline Truncated oleh Nisa' (2016) mengenai faktor-faktor yang memengaruhi jumlah kasus tuberkulosis di Jawa Timur di Jawa Timur didapatkan hasil bahwa presentase variabel gizi buruk masyarakat, presentase rumah tangga ber-PHBS, dan presentase tenaga kesehatan terlatih berpengaruh signifikan terhadap kasus tuberkulosis di Jawa Timur.

Berdasarkan referensi dari penelitian-penelitian pada kasus yang sama dan metode yang sama tersebut maka penelitian ini bertujuan untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus penyakit diare pada balita di Kota Surabaya. Berdasarkan data yang diperoleh dari Dinas Kesehatan Kota Surabaya tahun 2015, menunjukkan bahwa pola hubungan antara jumlah kasus penyakit diare pada balita dengan faktor-faktor yang diduga mempengaruhinya tidak memiliki pola tertentu, sehingga pada penelitian ini digunakan metode regresi nonparametrik. Menurut Eubank (1988) regresi nonparametrik adalah metode regresi yang digunakan ketika kurva regresinya tidak diketahui.



Fungsi yang digunakan dalam regresi nonparametrik pada penelitian ini adalah Spline Truncated. Spline Truncated merupakan potongan polinomial yang mempunyai sifat tersegmen, sehingga mempunyai fleksibilitas tinggi dan dapat menyesuaikan diri terhadap karakteristik lokal suatu data.

## **1.2 Rumusan Masalah**

Permasalahan yang dirumuskan dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik jumlah kasus penyakit diare pada balita di Surabaya ?
2. Bagaimana memodelkan jumlah kasus penyakit diare pada balita di Surabaya menggunakan Regresi Nonparametrik Spline Truncated ?
3. Bagaimana mengaplikasikan model yang diperoleh untuk analisis kebijakan pemerintah pada kasus penyakit diare balita?

## **1.3 Tujuan**

Tujuan dari penelitian ini adalah sebagai berikut.

1. Mengkaji karakteristik jumlah kasus penyakit diare pada balita di Surabaya berdasarkan variabel-variabel yang digunakan.
2. Memodelkan jumlah kasus penyakit diare pada balita di Surabaya menggunakan Regresi Nonparametrik Spline Truncated.
3. Mengaplikasikan model yang diperoleh untuk analisis kebijakan pemerintah pada kasus penyakit diare balita.

## **1.4 Manfaat**

Manfaat yang diharapkan melalui penelitian ini adalah sebagai berikut.

1. Penelitian ini diharapkan memberi informasi kepada pemerintah Kota Surabaya mengenai pemodelan Regresi Nonparametrik Spline pada kasus penyakit diare balita di Surabaya sehingga nantinya dapat dijadikan pertimbangan pembuatan kebijakan dalam pelaksanaan program-program

pemerintah di bidang kesehatan khususnya pada penyakit diare.

2. Memberikan informasi kepada masyarakat mengenai faktor-faktor paling signifikan yang mempengaruhi diare sehingga akan mengurangi angka kesakitan diare pada balita di Kota Surabaya
3. Hasil penelitian diharapkan dapat menjadi masukan dan acuan bagi penelitian selanjutnya.

### **1.5 Batasan Masalah**

Berdasarkan perumusan masalah yang telah dilakukan, maka batasan masalah pada penelitian ini adalah

1. Data yang digunakan dalam penelitian merupakan data sekunder pada tahun 2015 yang diambil dari Dinas Kesehatan Kota Surabaya.
2. Fungsi Spline yang digunakan adalah Spline Truncated Linear.
3. Titik knot yang digunakan adalah satu, dua, tiga, dan kombinasi knot.
4. Titik knot optimal dipilih dengan menggunakan metode *Generalized Cross Validation* (GCV).

## BAB II TINJAUAN PUSTAKA

### 2.1 Statistika Deskriptif

Statistika Deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna (Walpole, 1995). Statistika Deskriptif hanya memberikan gambaran data dan tidak dapat menarik suatu kesimpulan dari permasalahan yang ada. Statistika Deskriptif menyajikan data dalam bentuk ukuran pemusatan data, ukuran penyebaran data, diagram, tabel, grafik, serta kecenderungan suatu gugus data, sehingga data dapat dibaca secara ringkas dan menarik. Ukuran pemusatan data yang digunakan dalam penelitian ini adalah *mean* atau rata-rata dan median dengan rumus sebagai berikut:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.1)$$

Keterangan :

$\bar{x}$  : *mean*

$x_i$  : pengamatan ke- $i$ ,  $i=1,2,\dots,n$

$n$  : banyak pengamatan.

Ukuran penyebaran data yang digunakan adalah varians ( $s^2$ ). Berikut merupakan rumus varians:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (2.2)$$

### 2.2 Analisis Regresi

Analisis regresi adalah salah satu analisis Statistika yang bertujuan untuk menunjukkan hubungan matematis antara variabel respon dengan variabel prediktor (Gujarati, 2004). Pola hubungan matematis variabel tersebut pada umumnya dapat

digambarkan dengan *scatterplot* atau diagram pencar. Plot tersebut dapat menunjukkan apakah pola hubungan membentuk pola linier, kuadratik, kubik maupun tidak berbentuk pola atau acak. Terdapat tiga pendekatan dalam pencarian pola hubungan pada analisis regresi yaitu pendekatan regresi parametrik, pendekatan regresi nonparametrik, serta pendekatan regresi semiparametrik. Analisis regresi juga dapat digunakan untuk *forecasting* atau peramalan.

### 2.2.1 Regresi Nonparametrik

Regresi nonparametrik merupakan salah satu model regresi yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk kurva regresinya. Regresi nonparametrik merupakan model regresi yang sangat fleksibel dalam memodelkan pola data (Eubank, 1988). Model regresi nonparametrik secara umum dapat disajikan sebagai berikut:

$$y_i = \sum_{j=1}^p f_j(x_i) + \varepsilon_i, \quad i=1,2,3,\dots,n \quad (2.3)$$

dimana  $y_i$  merupakan variabel respon ke- $i$ ,  $f_j(x_i)$  merupakan fungsi regresi yang tidak diketahui bentuk kurva regresinya serta error random  $\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$ .

### 2.2.2 Regresi Nonparametrik Spline Truncated

Regresi nonparametrik Spline Truncated adalah regresi dimana fungsi atau kurva regresinya merupakan modifikasi dari fungsi polinomial. Fungsi Spline Truncated didapatkan dengan menjumlahkan antara fungsi polinomial dengan fungsi Truncated. Fungsi spline yang digunakan adalah fungsi spline truncated linear, karena lebih sederhana dan mudah diinterpretasikan. Sebagai contoh fungsi Spline Truncated berorde  $m$  dengan  $r$  titik titik knot. Sehingga bentuk kurva regresi  $f_j(x_i)$  secara lengkap adalah sebagai berikut.

$$f_j(x_i) = \sum_{h=0}^m \beta_{jh} x_j^h + \sum_{k=1}^r \beta_{j(k+m)} (x_j - K_{jk})_+^m \quad (2.4)$$

Apabila disubstitusikan ke persamaan 2.3 maka akan diperoleh persamaan model regresi nonparametrik Spline Truncated adalah sebagai berikut,

$$y_i = \sum_{j=1}^p \left( \sum_{h=0}^m \beta_{jh} x_j^h + \sum_{k=1}^r \beta_{j(k+m)} (x_j - K_{jk})_+^m + \varepsilon_i \right) \quad (2.5)$$

dimana,

$j$  = variabel prediktor,  $j = 1, 2, \dots, p$

$h$  = derajat polinomial,  $h = 0, 1, \dots, m$

$k$  = titik knot,  $k = 1, 2, \dots, r$

$i$  = banyaknya observasi,  $i = 1, 2, \dots, n$

$m = 1$  (karena fungsi spline truncated yang digunakan adalah fungsi linear)

Fungsi truncated  $(x_j - K_{jk})_+^m$  dapat dijabarkan dalam persamaan 2.6 berikut ini,

$$(x_j - K_{jk})_+^m = \begin{cases} (x_j - K_{jk})^m, & x_j \geq K_{jk} \\ 0 & , x_j < K_{jk} \end{cases} \quad (2.6)$$

Titik  $x_j = K_{jk}$  merupakan titik knot yang memperlihatkan pola perubahan dari fungsi pada sub interval yang berbeda dan nilai  $p$  merupakan derajat polinomial.

### 2.3 Estimasi Parameter

Salah satu metode yang dapat digunakan untuk mengestimasi parameter model regresi nonparametrik spline adalah *Ordinary Least Square* (OLS). Metode OLS mengestimasi parameter model regresi dengan meminimumkan jumlah kuadrat residual. Berikut merupakan bentuk penyajian matriks dari model regresi nonparametrik spline linear dengan  $r$  knot dan univariabel prediktor.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.7)$$

dimana

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_1 & (x_1 - K_1)_+^1 & \cdots & (x_1 - K_r)_+^1 \\ 1 & x_2 & (x_2 - K_1)_+^1 & \cdots & (x_2 - K_r)_+^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & (x_n - K_1)_+^1 & \cdots & (x_n - K_r)_+^1 \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Berdasarkan persamaan (2.7), persamaan residual dapat ditulis seperti Persamaan berikut,

$$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \quad (2.8)$$

Jumlah kuadrat residual dalam bentuk matriks dapat ditulis sebagai berikut,

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned}$$

Agar nilai  $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$  minimum, maka turunan pertama terhadap  $\boldsymbol{\beta}$  harus sama dengan nol.

$$\frac{\partial(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon})}{\partial\boldsymbol{\beta}} = 0$$

Persamaan terakhir memberikan:

$$\begin{aligned} -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= 0 \\ \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{y} \\ (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \end{aligned} \quad (2.9)$$

## 2.4 Pemilihan Titik Knot Optimal

Model regresi spline terbaik merupakan model yang memiliki titik knot optimal. Titik knot merupakan titik perpaduan

bersama dimana terdapat perubahan perilaku pola kurva pada interval yang berlainan (Budiantara, 2000). Metode yang digunakan untuk memilih titik knot optimal salah satunya yaitu dengan menggunakan metode *Generalized Cross Validation* (GCV). Metode GCV mempunyai sifat optimal asimtotik yang artinya dapat digunakan untuk sampel besar. Sifat optimal asimtotik ini tidak dimiliki oleh metode lain, misalnya *Cross Validation* (CV) (Wahba, 1990). Selain itu juga mempunyai sifat invarian terhadap transformasi yang artinya data dapat ditransformasi jika asumsi residual tidak terpenuhi. Model regresi spline terbaik diperoleh dari titik knot optimal dengan melihat nilai GCV terkecil. Metode GCV dapat dituliskan sebagai berikut (Eubank, 1988),

$$GCV(k) = \frac{MSE(k)}{[n^{-1}trace(\mathbf{I} - \mathbf{A}(k))]^2} \quad (2.10)$$

dimana  $\mathbf{I}$  adalah matriks identitas,  $n$  merupakan banyak pengamatan,  $k = (K_{1k}, K_{2k}, \dots, K_{pk})$ ,  $k = 1, 2, \dots, r$  merupakan titik-titik knot, dan

$$MSE(k) = n^{-1} \sum_{i=1}^n \left( y_i - \hat{f}(x_i) \right)^2$$

serta  $\mathbf{A}(k) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'$ .

## 2.5 Pengujian Parameter Model Regresi

Pengujian parameter model regresi digunakan untuk mengetahui apakah variabel prediktor memiliki pengaruh signifikan terhadap variabel respon. Pada regresi nonparametrik Spline Truncated ini, uji parameter model regresi dilakukan setelah mendapatkan model regresi dengan titik knot optimal berdasarkan nilai GCV yang paling minimum. Pengujian parameter dapat dibagi menjadi dua tahap yaitu pengujian secara serentak dan secara parsial.

### 2.5.1 Pengujian Secara Serentak

Pengujian parameter model secara serentak merupakan uji signifikansi seluruh parameter pada model regresi persamaan 2.5 dengan menggunakan uji  $F$  yang dapat dijabarkan pada tabel ANOVA pada tabel 2.1. berikut,

**Tabel 2.1** Analisis ragam (ANOVA) Uji Parameter

Sumber variasi	df	Sum of Square (SS)	Mean Square (MS)	$F_{hitung}$
Regresi	$p(m+r)$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	
Error	$n - p(m+r) - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{error}}{df_{error}}$	$\frac{MS_{regresi}}{MS_{error}}$
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	-	

dengan nilai  $p(m+r)$  merupakan banyak parameter dalam model regresi spline kecuali  $\beta_0$ .  $m$  adalah derajat polinomial dan  $r$  adalah titik knot optimal serta  $p$  adalah banyak variabel.

Hipotesis pada uji serentak sebagai berikut:

$$H_0 : \beta_{11} = \beta_{12} = \dots = \beta_{p(m+r)} = 0$$

$$H_1 : \text{minimal ada satu } \beta_{js} \neq 0; \quad j = 1, 2, \dots, p, s = 1, 2, \dots, m+r$$

Jika  $F_{hitung} > F_{\alpha(p(m+r); n-p(m+r)-1)}$  atau  $p\text{-value} < \alpha$  maka tolak  $H_0$  yang berarti bahwa minimal ada satu  $\beta_{js} \neq 0$ , atau minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon sehingga harus dilanjutkan pengujian secara parsial untuk mengetahui variabel-variabel prediktor mana yang berpengaruh signifikan.



### 2.5.2 Pengujian Secara Parsial atau Individu

Pengujian secara individu dilakukan untuk mengetahui apakah parameter secara individual mempunyai pengaruh yang signifikan terhadap variabel respon.

Hipotesis pada uji parsial adalah sebagai berikut:

$$H_0 : \beta_{js} = 0$$

$$H_1 : \beta_{js} \neq 0, \quad j = 1, 2, \dots, p, s = 1, 2, \dots, m + r$$

Pengujian secara individu dilakukan dengan menggunakan uji  $t$  (Draper dan Smith, 1992). Statistik uji yang digunakan adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_{js}}{SE(\hat{\beta}_{js})} \quad (2.11)$$

dengan  $SE(\hat{\beta}_{js})$  adalah *standart error*  $\hat{\beta}_{js}$  yang diperoleh dari akar elemen diagonal ke- $js$ ,  $j = 1, 2, \dots, p$  dan  $s = 1, 2, \dots, m + r$  dari matriks *variance-covariance*

$$\begin{aligned} Var(\hat{\beta}_{js}) &= \text{var} \left[ (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} \right] \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \text{var}(\mathbf{Y}) \left[ (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \right]' \\ &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' (\sigma^2 \mathbf{I}) \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

Jika  $|t_{hitung}| > t_{\alpha/2; (n-p(m+r)-1)}$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$  yang berarti variabel prediktor ke- $j$  tersebut berpengaruh signifikan terhadap variabel respon. Pengujian secara individu dilakukan satu per satu dari setiap variabel prediktor terhadap variabel respon.

### 2.6 Koefisien Determinasi ( $R^2$ )

Koefisien determinasi merupakan kuantitas yang dapat menjelaskan sumbangan variabel prediktor terhadap variabel

respon. Semakin tinggi nilai  $R^2$  yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabilitas variabel respon (Drapper dan Smith, 1992). Berikut ini merupakan rumus untuk menghitung  $R^2$ ,

$$R^2 = \frac{SS_{\text{Regresi}}}{SS_{\text{total}}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.12)$$

Selain itu, pemilihan model juga akan memperhatikan banyak parameter yang digunakan dalam model tersebut. Hal ini dijelaskan oleh prinsip parsimoni, dimana suatu model regresi yang baik adalah model regresi dengan banyak parameter yang sesedikit mungkin tetapi mempunyai  $R^2$  yang cukup tinggi.

## 2.7 Pengujian Asumsi Residual Model Regresi

Pengujian asumsi residual (*goodness of fit*) model regresi paling populer karena mudah digunakan. Residual yang dihasilkan harus memenuhi asumsi. Terdapat tiga asumsi yang harus dipenuhi yaitu identik, independen, dan berdistribusi normal.

### 2.7.1 Asumsi Identik

Asumsi identik (homoskedastisitas) berarti bahwa varians pada residual sama atau identik. Kebalikanannya adalah kasus heteroskedastisitas, yaitu jika kondisi varians *residual* tidak identik (Gujarati, 2009).

$$\text{var}(y_i) = \text{var}(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n$$

Uji identik dapat menggunakan uji Glejser. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah seperti Persamaan (2.13) berikut,

$$F_{hitung} = \frac{\frac{\sum_{i=1}^n (|\hat{\varepsilon}_i| - |\bar{\varepsilon}|)^2}{p(m+r)}}{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)^2}{n - p(m+r) - 1}} \quad (2.13)$$

dimana nilai  $p(m+r)$  adalah banyaknya parameter model Glejser dan untuk model regresi nonparametrik Spine Truncated seperti persamaan 2.5

Jika  $F_{hitung} > F_{\alpha(p(m+r), n-p(m+r)-1)}$  atau  $p\text{-value} < \alpha$  maka tolak  $H_0$  yang berarti bahwa tidak terindikasi terdapat kasus homoskedastisitas dan sebaliknya jika  $F_{hitung} < F_{\alpha(p(m+r), n-p(m+r)-1)}$  atau  $p\text{-value} > \alpha$  maka gagal tolak  $H_0$  yang berarti bahwa terindikasi terdapat kasus homoskedastisitas sehingga asumsi identik terpenuhi.

### 2.7.2 Asumsi Independen (Uji Autokorelasi)

Asumsi independen merupakan asumsi dari model regresi yang mengharuskan tidak terdapat korelasi antar residual. Ada beberapa uji yang dapat digunakan untuk mendeteksi kasus autokorelasi salah satunya adalah dengan melakukan plot antara residual dengan *observation order*. Apabila plot secara visual menunjukkan acak maka asumsi independen dapat terpenuhi.

Selain itu untuk membuktikan asumsi independen dapat menggunakan *Run Test*. Untuk hipotesis yang digunakan dalam uji Run test adalah sebagai berikut:

$H_0 : \rho = 0$  (residual memenuhi asumsi independen)

$H_1 : \rho \neq 0$  (residual tidak memenuhi asumsi independen)

Statistik Uji

$$Z = \frac{r - \{[(2n_1n_2) / (n_1 + n_2)] + 1\}}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}} \quad (14)$$

Daerah penolakan, tolak  $H_0$  apabila nilai  $Z < -Z_{\alpha/2}$  atau  $Z > Z_{\alpha/2}$  yang menandakan residual tidak memenuhi asumsi independen dan juga sebaliknya. Selain itu apabila nilai  $P_{value}$  pada *Run test* kurang dari  $\alpha$  (0.05) maka tolak  $H_0$  yang berarti asumsi independen tidak terpenuhi dan sebaliknya.

### 2.7.3 Asumsi Berdistribusi Normal

Uji *Kolmogorov-Smirnov* bertujuan untuk mengetahui apakah suatu data telah mengikuti suatu distribusi tertentu. Hipotesis yang digunakan :

$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$  atau residual berdistribusi normal

$H_0 : F_n(\varepsilon) \neq F_0(\varepsilon)$  atau residual tidak berdistribusi normal

Statistik uji :

$$D = \sup_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (2.14)$$

Tolak  $H_0$  atau residual tidak berdistribusi normal apabila  $D > D_\alpha$  dimana,

$D_\alpha$  adalah nilai kritis untuk uji *Kolmogorov Smirnov* satu sampel, diperoleh dari tabel *Kolmogorov Smirnov* satu sampel,  $F_n(\varepsilon)$  adalah nilai peluang kumulatif (fungsi distribusi kumulatif) berdasarkan data sampel,  $F_0(\varepsilon)$  adalah nilai peluang kumulatif (fungsi distribusi kumulatif) dibawah  $H_0$ .

## 2.8 Skenario Model

Skenario model dilakukan untuk melakukan prediksi nilai variabel respon dengan menggunakan simulasi nilai-nilai variabel prediktor pada model regresi nonparametrik Spline Truncated yang terpilih. Skenario model terdiri dari skenario optimis, *middle*, dan pesimis. Setiap skenario dilakukan sebanyak tiga kali. Penentuan range setiap skenario model menggunakan range kuartil 1 dan kuartil 3.

## 2.9 Diare

Diare didefinisikan sebagai bertambahnya defekasi (buang air besar) lebih dari biasanya/lebih dari tiga kali sehari, disertai dengan perubahan konsisten tinja (menjadi cair) dengan atau tanpa darah (WHO, 2010). Secara klinik dibedakan tiga macam sindroma diare yaitu diare cair akut, disentri, dan diare persisten. Diare akut diberi batasan sebagai meningkatnya kekerapan, bertambah cairan, atau bertambah banyaknya tinja yang dikeluarkan, akan tetapi hal itu sangat relatif terhadap kebiasaan yang ada pada penderita dan berlangsung tidak lebih dari satu minggu. Apabila diare berlangsung antara satu sampai dua minggu maka dikatakan diare yang berkepanjangan (Soegijanto, 2002). Beberapa perilaku yang dapat meningkatkan risiko terjadinya diare pada balita, yaitu (Depkes RI, 2007) :

1. Tidak memberikan ASI secara penuh 4-6 bulan pertama pada kehidupan. Pada balita yang tidak diberi ASI resiko menderita diare lebih besar daripada balita yang diberi ASI penuh, dan kemungkinan menderita dehidrasi berat lebih besar.
2. Menggunakan botol susu, penggunaan botol ini memudahkan pencemaran oleh kuman karena botol susah dibersihkan. Penggunaan botol yang tidak bersih atau sudah dipakai selama berjam-jam dibiarkan dilingkungan yang panas, sering menyebabkan infeksi usus yang parah karena botol dapat tercemar oleh kuman-kuman/bakteri penyebab diare.
3. Menyimpan makanan masak pada suhu kamar, bila makanan disimpan beberapa jam pada suhu kamar, makanan akan tercemar dan kuman akan berkembang biak.
4. Menggunakan air minum yang tercemar.
5. Tidak mencuci tangan sesudah buang air besar dan sesudah membuang tinja anak atau sebelum makan dan menyuapi anak
6. Tidak membuang tinja dengan benar, seringnya beranggapan bahwa tinja tidak berbahaya, padahal sesungguhnya

mengandung virus atau bakteri dalam jumlah besar. Selain itu tinja binatang juga dapat menyebabkan infeksi pada manusia

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari, Badan Pusat Statistik(BPS) Kota Surabaya, Dinas Kesehatan Kota Surabaya, Profil Kesehatan Provinsi Kota Surabaya tahun 2015. Unit penelitian yang digunakan merupakan 31 Kecamatan di Kota Surabaya.

#### **3.2 Variabel Penelitian**

Variabel respon yang digunakan dalam penelitian ini adalah jumlah kasus penyakit diare yang ditangani di Kota Surabaya, sedangkan variabel prediktor yang diduga berpengaruh terhadap kasus penyakit diare diperoleh dari Departemen Kesehatan RI serta penelitian-penelitian sebelumnya yang diuraikan dalam Tabel 3.1 berikut,

**Tabel 3.1** Variabel Penelitian

<b>Variabel</b>	<b>Keterangan</b>
$y$	Presentase Jumlah Kasus Penyakit diare pada balita (%)
$x_1$	Kepadatan Penduduk (ribu jiwa/km <sup>2</sup> )
$x_2$	Persentase rumah tangga yang berperilaku hidup bersih dan sehat (PHBS) (%)
$x_3$	Presentase penduduk dengan akses sanitasi layak (jamban sehat) (%)
$x_4$	Presentase bayi yang diberi ASI eksklusif pada usia 0-6 bulan (%)

Berikut ini merupakan keterangan dari variabel penelitian menurut Dinas Kesehatan Kota Surabaya.

- a. Variabel  $y$  merupakan presentase jumlah kasus diare pada balita. Nilai tersebut didapatkan dengan rumus berikut,

$$\frac{\text{Jumlah Kasus Diare di wilayah dan waktu tertentu}}{\text{Jumlah Balita di wilayah dan kurun waktu yang sama}} \times 100\%$$

Variabel  $x_1$  merupakan kepadatan penduduk. Kepadatan penduduk adalah banyaknya jumlah penduduk untuk setiap kilometer persegi luas wilayah. Dalam analisis ini data yang digunakan dalam satuan ribu jiwa per km<sup>2</sup>.

- b. Variabel  $x_2$  merupakan persentase rumah tangga yang ber-PHBS (%). Nilai tersebut didapatkan dari rumus berikut,

$$\frac{\text{Jumlah Rumah Tangga berperilaku hidup bersih sehat}}{\text{Jumlah Rumah Tangga yang dipantau}} \times 100\%$$

dimana rumah tangga yang berperilaku hidup bersih sehat (PHBS) adalah rumah tangga yang seluruh anggotanya berperilaku hidup bersih dan sehat, yang meliputi 10 indikator, yaitu pertolongan persalinan oleh tenaga kesehatan, bayi diberi ASI eksklusif, balita ditimbang setiap bulan, menggunakan air bersih, mencuci tangan dengan air bersih dan sabun, menggunakan jamban sehat, memberantas jentik di rumah seminggu sekali, makan sayur dan buah setiap hari, melakukan aktivitas fisik setiap hari, dan tidak merokok didalam rumah.

- c. Variabel  $x_3$  merupakan presentase penduduk dengan akses sanitasi layak (jamban sehat). Nilai tersebut didapatkan dari rumus berikut,

$$\frac{\text{Jumlah Penduduk dengan akses sanitasi layak}}{\text{Jumlah Penduduk di wilayah dan periode yang sama}} \times 100\%$$

Fasilitas sanitasi yang layak (jamban sehat) adalah fasilitas sanitasi yang memenuhi syarat kesehatan antara lain dilengkapi dengan leher angsa, tanki septik/Sistem Pengolahan Air Limbah (SPAL), yang digunakan sendiri atau bersama.



- d. Variabel  $x_4$  merupakan presentase bayi yang diberi ASI eksklusif pada usia 0-6 bulan Nilai tersebut didapatkan dari rumus berikut,

$$\frac{\text{Jumlah Bayi dengan ASI eksklusif 0-6 bulan}}{\text{Jumlah Bayi di wilayah dan periode yang sama}} \times 100\%$$

### 3.3 Struktur Data

Struktur data yang terdiri dari variabel respon dan empat variabel prediktor yang digunakan dalam penelitian ini ditunjukkan pada Tabel 3.2

**Tabel 3.2** Struktur Data

Kecamatan	$y$	$x_1$	$x_2$	$x_3$	$x_4$
1	$y_1$	$x_{1(1)}$	$x_{2(1)}$	$x_{3(1)}$	$x_{4(1)}$
2	$y_2$	$x_{1(2)}$	$x_{2(2)}$	$x_{3(2)}$	$x_{4(2)}$
3	$y_3$	$x_{1(3)}$	$x_{2(3)}$	$x_{3(3)}$	$x_{4(3)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
31	$y_{31}$	$x_{1(31)}$	$x_{2(31)}$	$x_{3(31)}$	$x_{4(31)}$

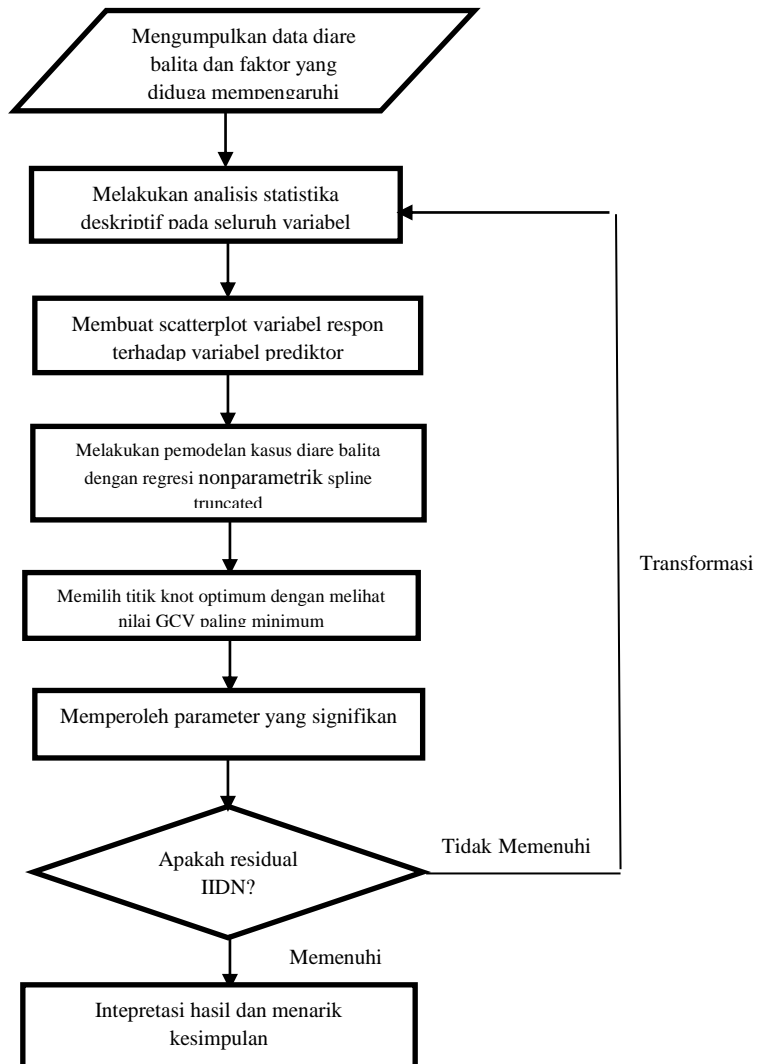
### 3.4 Langkah Analisis

Langkah-langkah analisis dalam penelitian adalah sebagai berikut.

1. Mendeskripsikan kasus kesakitan diare di Kota Surabaya dan faktor-faktor yang mempengaruhinya.
2. Membuat *scatter plot* antara variabel respon dengan masing-masing variabel prediktor untuk mengetahui pola hubungan yang terjadi.
3. Menentukan komponen nonparametrik secara visual dengan menggunakan *scatterplot*
4. Memodelkan variabel respon dan variabel prediktor menggunakan model regresi nonparametrik spline truncated linear dengan satu titik knot.
5. Memodelkan variabel respon dan variabel prediktor menggunakan model regresi nonparametrik spline dengan dua titik knot.

6. Memodelkan variabel respon dan variabel prediktor menggunakan model regresi nonparametrik spline dengan tiga titik knot.
7. Memodelkan variabel respon dan variabel prediktor menggunakan model regresi nonparametrik spline truncated linear dengan kombinasi titik knot.
8. Memilih titik knot optimal berdasarkan nilai GCV yang paling minimum.
9. Mendapatkan model regresi spline terbaik dengan titik knot optimal.
10. Melakukan uji signifikansi parameter secara serentak dan parsial.
11. Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN) dari model regresi spline.
12. Membuat interpretasi model dan menarik kesimpulan berupa rekomendasi kebijakan.

### 3.5 Diagram Alir Penelitian



**Gambar 3.1** Diagram Alir Penelitian

*(Halaman ini sengaja dikosongkan)*

## BAB IV

### ANALISIS DAN PEMBAHASAN

Pada bab ini akan dijelaskan mengenai gambaran umum dari jumlah penderita diare pada balita di Kota Surabaya tahun 2015 beserta dengan faktor-faktor yang diduga memengaruhinya. Gambaran umum pada kasus ini akan dijabarkan dengan menggunakan statistika deskriptif yang terdiri dari nilai rata-rata, varians, serta nilai minimum dan maksimum. Untuk pemodelan jumlah penderita diare pada balita di Kota Surabaya beserta faktor-faktor yang diduga mempengaruhi dapat dianalisa terlebih dahulu dengan menggunakan *scatterplot* untuk melihat apakah variabel tersebut termasuk komponen parametrik ataupun nonparametrik. Karena seluruh variabel prediktor mengikuti pola nonparametrik maka pemodelan yang dilakukan adalah dengan menggunakan metode regresi nonparametrik Spline Truncated dengan satu knot, dua knot, tiga knot, dan kombinasi knot. Untuk fungsi Spline Truncated yang digunakan pada penelitian ini adalah fungsi Linier. Sedangkan untuk pemodelannya akan dilakukan dengan pemilihan titik knot optimum dari setiap knot mulai dengan menggunakan satu titik knot hingga kombinasi titik knot dengan memperhatikan nilai GCV terkecil. Untuk analisis lebih lanjut akan dijelaskan pada sub bab berikutnya.

#### **4.1 Gambaran Umum Jumlah Penderita Diare Balita di Kota Surabaya dan Faktor yang Diduga Mempengaruhi**

Langkah pertama sebelum melakukan analisis adalah dengan mengetahui pola dari data yang akan dianalisis. Pola dapat disajikan dengan gambaran umum data menggunakan statistika deskriptif seperti rata-rata, varians, serta nilai minimum dan nilai maksimum dari data tersebut. Seluruh variabel mulai variabel respon hingga keempat variabel prediktor yang diduga mempengaruhi dilakukan perhitungan rata-rata, varians hingga nilai minimum dan maksimumnya. Hasil perhitungan statistika deskriptif dapat disajikan dalam Tabel 4.1 seperti berikut.

**Tabel 4.1** Gambaran umum Kasus Diare Balita di Surabaya dan Faktor-Faktor yang Diduga Mempengaruhinya

Variabel	Rata-rata	Varians	Minimum	Maksimum
$y$	7,849	28,169	2,314	28,828
$x_1$	11,47	52,54	2,21	31,54
$x_2$	72,77	160,56	44,33	98,66
$x_3$	97,981	4,881	91,546	100
$x_4$	64,59	136,22	36,54	88,26

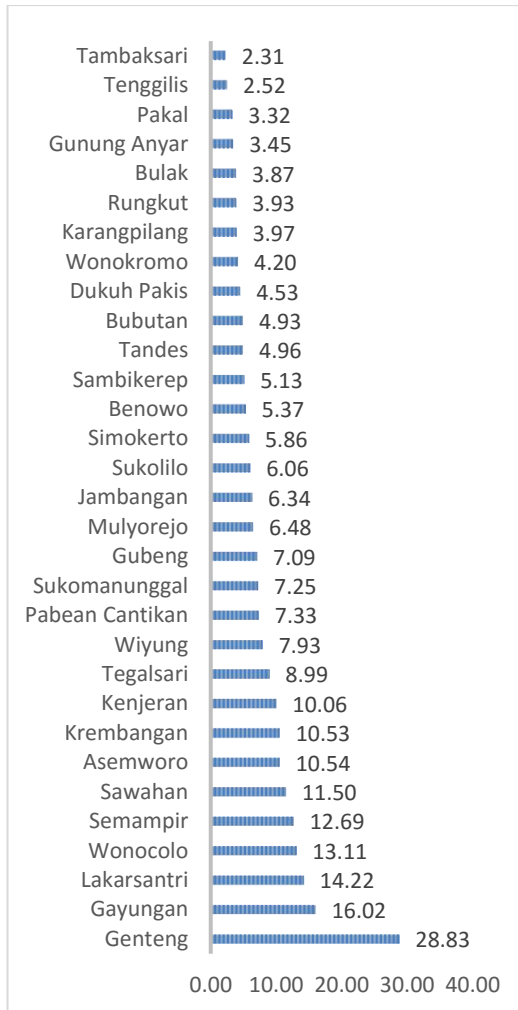
Berdasarkan hasil perhitungan yang ditunjukkan dari Tabel 4.1 menunjukkan informasi sebagai berikut.

- Variabel  $y$  yaitu presentase kasus penyakit diare pada balita di Kota Surabaya pada tahun 2015. Nilai tersebut menunjukkan perbandingan antara jumlah kasus diare yang terjadi pada balita di tahun 2015 terhadap seluruh balita yang ada di daerah tersebut pada periode yang sama. Nilai rata-rata dari kasus terjadinya penyakit diare balita tahun 2015 adalah sebesar 7,849%, nilai varians sebesar 28,169, nilai minimum sebesar 2,314% dan nilai maksimum sebesar 28,828%. Dilihat dari nilai varians yang cukup besar menandakan bahwa presentase kasus diare balita di Kota Surabaya tidak merata. Presentase kasus diare terendah yaitu di Kecamatan Tambaksari dan tertinggi yaitu di Kecamatan Genteng.
- Variabel  $x_1$  yaitu kepadatan penduduk di Kota Surabaya tahun 2015. Nilai tersebut diperoleh dari perbandingan antara jumlah penduduk pada tiap kecamatan di Kota Surabaya tahun 2015 dibandingkan dengan luas wilayah pada kecamatan tersebut dalam periode waktu yang sama. Pada penelitian ini kepadatan penduduk dihitung berdasarkan satuan ribu jiwa per  $\text{km}^2$ . Nilai rata-rata sebesar 11,47 ribu jiwa/ $\text{km}^2$ , nilai varians sebesar 52,54, nilai minimum sebesar 2,21 ribu jiwa/ $\text{km}^2$  dan nilai maksimum sebesar 31,54 ribu jiwa/ $\text{km}^2$ . Dilihat dari nilai varians yang sangat besar menandakan bahwa kepadatan penduduk di Kota Surabaya tidak merata. Kepadatan

penduduk terendah yaitu di Kecamatan Pakal dan tertinggi yaitu di Kecamatan Simokerto.

- c. Variabel  $x_2$  yaitu presentase rumah tangga berperilaku hidup bersih sehat (PHBS) di Kota Surabaya tahun 2015. Berdasarkan data tersebut nilai rata-rata rumah tangga berperilaku hidup bersih sehat adalah sebesar 72,77%, nilai varians sebesar 160,56, nilai minimum sebesar 44,33% dan nilai maksimum sebesar 98,66%. Dilihat dari nilai varians yang cukup besar menandakan bahwa presentase rumah tangga ber-PHBS di Kota Surabaya tidak merata. Presentase rumah tangga ber-PHBS di Kota Surabaya terendah yaitu di Kecamatan Semampir dan tertinggi yaitu di Kecamatan Tenggilis.
- d. Variabel  $x_3$  yaitu presentase penduduk dengan akses jamban sehat di Kota Surabaya tahun 2015 memiliki nilai rata-rata sebesar 97,981%, nilai varians sebesar 4,881, nilai minimum sebesar 91,546% dan nilai maksimum sebesar 100%. Dilihat dari nilai varians yang cukup kecil menandakan bahwa penduduk dengan akses jamban sehat di Kota Surabaya cukup merata. Penduduk dengan akses jamban sehat terendah yaitu di Kecamatan Asemrowo dan tertinggi yaitu di Kecamatan Dukuh Pakis dan Sambikerep.
- e. Variabel  $x_4$  yaitu presentase balita dengan ASI eksklusif 0-6 bulan di Kota Surabaya tahun 2015 memiliki nilai rata-rata sebesar 64,59%, nilai varians sebesar 136,22, nilai minimum sebesar 36,54% dan nilai maksimum sebesar 88,26%. Dilihat dari nilai varians yang cukup besar menandakan bahwa balita dengan ASI eksklusif 0-6 bulan di Kota Surabaya kurang merata. balita dengan ASI eksklusif 0-6 bulan terendah yaitu di Kecamatan Tandes dan tertinggi yaitu di Kecamatan Pakal.

Jika disajikan dalam bentuk grafik maka akan terlihat presentase kasus diare balita menurut 31 Kecamatan di Kota Surabaya tahun 2015 mulai dari jumlah kasus diare balita paling rendah hingga kasus diare balita dengan jumlah paling tinggi. Berikut ini adalah presentase kasus diare balita di Kota Surabaya.



**Gambar 4.1** Presentase Kasus Diare Balita tiap Kecamatan di Kota Surabaya Tahun 2015

Gambar 4.1 menunjukkan bahwa kecamatan di Kota Surabaya yang memiliki presentase diare balita tertinggi adalah Kecamatan Genteng sebesar 28,83% dan yang terendah adalah Kecamatan Tambaksari sebesar 2,31%. Terdapat sembilan

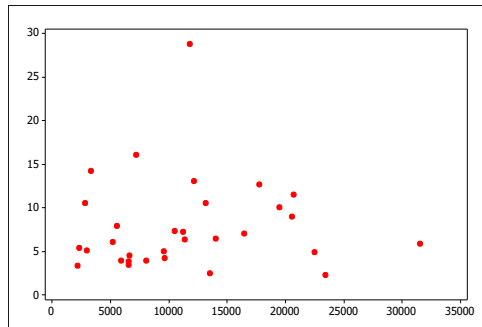


kecamatan yang memiliki presentase kasus diare balita lebih dari 10%.

#### 4.2 *Scatterplot* Jumlah Penderita Diare Balita dengan Faktor-faktor yang Diduga Mempengaruhi

Setelah gambaran umum dari data penderita diare balita dan faktor pengaruhnya diketahui maka langkah selanjutnya adalah dengan melihat *scatterplot* dari variabel respon dengan masing-masing variabel prediktor untuk melihat pola hubungan antar variabel respon terhadap seluruh variabel prediktor yang dilakukan secara visual.

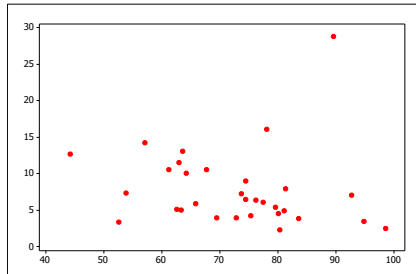
- a. *Scatterplot* antara Jumlah Penderita Diare dengan Kepadatan Penduduk.



**Gambar 4.2** *Scatterplot* Antara Jumlah Penderita Diare Balita dengan Kepadatan Penduduk

Berdasarkan Gambar 4.2 menunjukkan bahwa plot antara variabel jumlah penderita diare balita dengan kepadatan penduduk tidak mengikuti atau tidak membentuk pola tertentu. Sehingga variabel kepadatan penduduk termasuk kedalam komponen nonparametrik. Secara universal, jika kepadatan penduduk meningkat maka jumlah penderita diare pada balita cenderung meningkat juga.

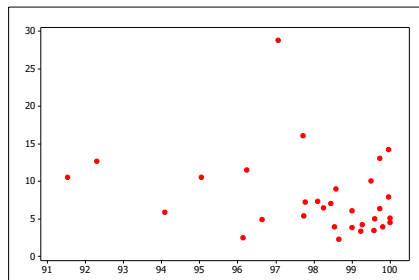
b. *Scatterplot* antara Jumlah Penderita Diare dengan Rumah Tangga ber-PHBS



**Gambar 4.3** *Scatterplot* Antara Jumlah Penderita Diare Balita dengan Rumah Tangga ber-PHBS

Berdasarkan Gambar 4.3 menunjukkan bahwa plot antara variabel jumlah penderita diare balita dengan variabel rumah tangga ber-PHBS tidak mengikuti atau tidak membentuk pola tertentu. Sehingga variabel rumah tangga ber-PHBS termasuk kedalam komponen nonparametrik. Secara universal, jika rumah tangga ber-PHBS meningkat maka jumlah penderita diare pada balita cenderung menurun atau berkurang.

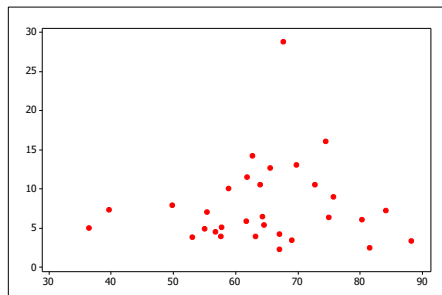
c. *Scatterplot* antara Jumlah Penderita Diare dengan Penduduk dengan Akses Jamban Sehat



**Gambar 4.4** *Scatterplot* Antara Jumlah Penderita Diare Balita dengan Penduduk dengan Akses jamban Sehat

Berdasarkan Gambar 4.4 menunjukkan bahwa plot antara variabel jumlah penderita diare balita dengan variabel penduduk dengan akses jamban sehat tidak mengikuti atau tidak membentuk pola tertentu. Sehingga variabel penduduk dengan akses jamban sehat termasuk kedalam komponen nonparametrik. Secara universal, jika penduduk dengan akses jamban sehat meningkat maka jumlah penderita diare pada balita cenderung menurun atau berkurang.

d. *Scatterplot* antara Jumlah Penderita Diare dengan Bayi Mendapatkan ASI Eksklusif 0-6 bulan



**Gambar 4.5** *Scatterplot* Antara Jumlah Penderita Diare Balita dengan Bayi Mendapatkan ASI Eksklusif 0-6 bulan

Berdasarkan Gambar 4.5 menunjukkan bahwa plot antara variabel jumlah penderita diare balita dengan variabel bayi mendapatkan ASI eksklusif 0-6 bulan tidak mengikuti atau tidak membentuk pola tertentu. Sehingga variabel bayi mendapatkan ASI eksklusif 0-6 bulan termasuk kedalam komponen nonparametrik. Secara universal, jika bayi mendapatkan ASI eksklusif 0-6 bulan meningkat maka jumlah penderita diare pada balita cenderung menurun atau berkurang.

### 4.3 Pemodelan Jumlah Penderita Diare Balita Menggunakan Regresi Nonparametrik Spline Truncated

Pemodelan jumlah penderita diare pada balita di Kota Surabaya meliputi jumlah penderita diare pada balita di Kota Surabaya sebagai variabel respon, dan faktor-faktor lain yang diduga mempengaruhi sebagai variabel prediktor menggunakan metode regresi nonparametrik Spline Truncated. Setelah itu dilakukan estimasi model regresi nonparametrik Spline Truncated dengan menggunakan satu knot, dua knot, tiga knot, dan kombinasi knot.

### 4.4 Pemilihan Titik Knot Optimum

Dalam pendekatan regresi nonparametrik Spline Truncated, dikenal adanya titik knot. Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku data. Dalam pemilihan titik knot optimum pada variabel-variabel yang diduga mempengaruhi jumlah penderita diare balita di Kota Surabaya dilakukan dengan menggunakan metode *Generalized Cross Validation* (GCV). Titik knot optimal diperoleh dari nilai GCV yang paling minimum.

#### 4.4.1 Regresi Nonparametrik Spline Truncated dengan Satu Knot

Pada model regresi nonparametrik Spline Truncated dengan satu knot, model yang terbentuk adalah sebagai berikut

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+ + \hat{\beta}_{21}x_2 + \hat{\beta}_{22}(x_2 - K_2)_+ + \hat{\beta}_{31}x_3 + \hat{\beta}_{32}(x_3 - K_3)_+ + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_4)_+$$

Estimasi parameter pada regresi nonparametrik Spline dengan satu knot menghasilkan beberapa iterasi, Berikut ditampilkan 10 hasil iterasi yang juga mencakup iterasi dengan nilai GCV terkecil,

**Tabel 4.2** Pemilihan Titik Knot Optimum dengan Satu Titik Knot

GCV	$x_1$	$x_2$	$x_3$	$x_4$
43,29	2,813	45,439	91,719	37,596
45,55	8,199	55,418	93,271	47,095
44,28	14,183	66,506	94,997	57,650
42,74	15,978	69,832	95,514	60,817
42,21	17,175	72,049	95,859	62,928
40,48	18,970	75,376	96,377	66,094
<b>35,56</b>	<b>21,364</b>	<b>79,811</b>	<b>97,067</b>	<b>70,316</b>
38,91	23,758	84,246	97,757	74,538
42,98	26,151	88,681	98,447	78,760
43,08	30,939	97,551	99,827	87,204

Berdasarkan Tabel 4.2, nilai GCV paling minimum adalah 35,56 yang terdapat pada iterasi ke-32. Nilai titik knot untuk variabel kepadatan penduduk ( $x_1$ ) adalah 21,364, variabel presentase rumah tangga ber-PHBS ( $x_2$ ) adalah 79,811 variabel presentase penduduk dengan sanitasi jamban sehat ( $x_3$ ) adalah 97,067 dan variabel presentase bayi yang diberi ASI eksklusif ( $x_4$ ) adalah 70,316.

Selanjutnya, hasil dari GCV dengan menggunakan satu titik knot akan dibandingkan dengan hasil dari GCV dengan menggunakan dua titik knot, dan tiga titik knot. Perbandingan hasil GCV tersebut dilakukan untuk memperoleh nilai GCV yang paling minimum dan diharapkan dapat menghasilkan model *Spline* terbaik.

#### **4.4.2 Regresi Nonparametrik Spline Truncated dengan Dua Knot**

Pada model regresi nonparametrik Spline Truncated dengan dua knot, model yang terbentuk adalah sebagai berikut

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+ + \hat{\beta}_{13}(x_1 - K_2)_+ + \hat{\beta}_{21}x_2 + \hat{\beta}_{22}(x_2 - K_3)_+ + \hat{\beta}_{23}(x_2 - K_4)_+ + \hat{\beta}_{31}x_3 + \hat{\beta}_{32}(x_3 - K_5)_+ + \hat{\beta}_{33}(x_3 - K_6)_+ + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_7)_+ + \hat{\beta}_{43}(x_4 - K_8)_+$$

Estimasi parameter pada regresi nonparametrik Spline dengan dua knot menghasilkan beberapa iterasi. Pemilihan titik knot optimum dengan melihat nilai GCV paling minimum. Berikut ditampilkan 10 hasil iterasi yang juga mencakup iterasi dengan nilai GCV terkecil,

**Tabel 4.3** Pemilihan Titik Knot Optimum dengan Dua Titik Knot

GCV	$x_1$	$x_2$	$x_3$	$x_4$
	2,813	45,439	91,719	37,594
50,06	3,411	46,548	91,891	38,649
	3,411	46,548	91,891	38,649
58,24	9,994	58,744	93,789	50,260
	4,608	48,765	92,236	40,760
53,52	5,207	49,874	92,409	41,816
	5,805	50,983	92,582	42,871
54,03	15,978	69,832	95,514	60,814
	7,600	54,309	93,099	46,038
56,92	14,183	66,506	94,997	57,648
	10,592	59,853	93,962	51,315
50,30	20,765	78,702	96,895	69,258
	14,781	67,614	95,169	58,703
51,22	23,757	84,246	97,757	74,535
	<b>21,364</b>	<b>79,811</b>	<b>97,067</b>	<b>70,316</b>
<b>13,66</b>	<b>26,750</b>	<b>89,790</b>	<b>98,620</b>	<b>79,816</b>

**Tabel 4.3** Pemilihan Titik Knot Optimum dengan Dua Titik Knot (Lanjutan)

GCV	$x_1$	$x_2$	$x_3$	$x_4$
	29,143	94,225	99,310	84,034
47,23	30,340	96,442	99,655	86,145
	30,340	96,442	99,655	86,145
47,25	30,938	97,551	99,827	87,201

Berdasarkan Tabel 4.2, nilai GCV paling minimum adalah 13,66 yang terdapat pada iterasi ke-1001. Nilai titik knot untuk variabel kepadatan penduduk ( $x_1$ ) adalah  $K_1= 21,364$  dan  $K_2= 26,750$ , variabel presentase rumah tangga ber-PHBS ( $x_2$ ) adalah  $K_3= 79,811$  dan  $K_4= 89,790$  variabel presentase penduduk dengan sanitasi jamban sehat ( $x_3$ ) adalah  $K_5= 97,067$  dan  $K_6= 98,620$  dan variabel presentase bayi yang diberi ASI eksklusif ( $x_4$ ) adalah  $K_7= 70,316$  dan  $K_8= 79,816$ .

Selanjutnya, hasil dari GCV dengan menggunakan satu titik knot akan dibandingkan dengan hasil dari GCV dengan menggunakan dua titik knot, dan tiga titik knot. Perbandingan hasil GCV tersebut dilakukan untuk memperoleh nilai GCV yang paling minimum untuk menghasilkan model *Spline* terbaik.

#### 4.4.3 Regresi Nonparametrik Spline Truncated dengan Tiga Knot

Pada model regresi nonparametrik Spline Truncated dengan tiga knot, model yang terbentuk adalah sebagai berikut

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+ + \hat{\beta}_{13}(x_1 - K_2)_+ + \hat{\beta}_{14}(x_1 - K_3)_+ + \hat{\beta}_{21}x_2 + \hat{\beta}_{22}(x_2 - K_4)_+ + \hat{\beta}_{23}(x_2 - K_5)_+ + \hat{\beta}_{24}(x_2 - K_6)_+ + \hat{\beta}_{31}x_3 + \hat{\beta}_{32}(x_3 - K_7)_+ + \hat{\beta}_{33}(x_3 - K_8)_+ + \hat{\beta}_{34}(x_3 - K_9)_+ + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_{10})_+ + \hat{\beta}_{43}(x_4 - K_{11})_+ + \hat{\beta}_{44}(x_4 - K_{12})_+$$

Estimasi parameter pada regresi nonparametrik Spline dengan tiga knot menghasilkan beberapa iterasi Berikut

ditampilkan 10 hasil iterasi yang juga mencakup iterasi dengan nilai GCV terkecil.

**Tabel 4.4** Pemilihan Titik Knot Optimum dengan Tiga Titik Knot

GCV	$x_1$	$x_2$	$x_3$	$x_4$
52,81	2,813	45,439	91,719	37,594
	3,411	46,548	91,891	38,649
	4,010	47,656	92,064	39,705
	3,411	46,548	91,891	38,649
	7,002	53,2	92,927	44,982
71,41	3,034	96,442	99,655	86,145
<b>3,412    46,548    91,891    38,651</b>				
<b>19,569    76,484    96,549    67,150</b>				
<b>12,73</b>	<b>26,750</b>	<b>89,790</b>	<b>98,620</b>	<b>79,816</b>
74,37	4,608	48,765	92,236	40,76
	5,805	50,983	92,582	42,871
	24,954	86,463	98,102	76,646
	5,207	49,874	92,409	41,816
	25,552	87,572	98,275	77,702
24,01	26,749	89,79	98,62	79,813
60,47	8,199	55,418	93,272	47,093
	14,781	67,614	95,169	58,703
	29,741	95,334	99,482	85,09
	11,191	60,962	94,134	52,37
	29,741	95,334	99,482	85,09
60,43	30,340	96,442	99,655	86,145



**Tabel 4.4** Pemilihan Titik Knot Optimum dengan Tiga Titik Knot (Lanjutan)

GCV	$x_1$	$x_2$	$x_3$	$x_4$
	27,946	92,007	98,965	81,923
	28,544	93,116	99,137	82,979
64,08	29,143	94,225	99,31	84,034
	29,143	94,225	99,31	84,034
	30,340	96,442	99,655	86,145
59,56	30,938	97,551	99,827	87,201
	29,741	95,334	99,482	85,09
	30,340	96,442	99,655	86,145
52,06	30,938	97,551	99,827	87,201

Berdasarkan Tabel 4.4, nilai GCV paling minimum adalah 12,73 yang terdapat pada iterasi ke-1938. Nilai titik knot untuk masing-masing variabel adalah sebagai berikut.

$$\begin{aligned}
 K_1 &= 3,412; & K_2 &= 19,569; & K_3 &= 26,750; \\
 K_4 &= 46,548; & K_5 &= 76,484; & K_6 &= 89,790; \\
 K_7 &= 91,891; & K_8 &= 96,549; & K_9 &= 98,620; \\
 K_{10} &= 38,651 & K_{11} &= 67,150; & K_{12} &= 79,816;
 \end{aligned}$$

#### 4.4.4 Regresi Nonparametrik Spline Truncated dengan Kombinasi Knot

Setelah melakukan pemodelan regresi nonparametrik spline truncated satu, dua, dan tiga knot dilanjutkan dengan pemodelan menggunakan kombinasi knot. Estimasi parameter pada regresi nonparametrik Spline dengan kombinasi knot menghasilkan beberapa iterasi, Berikut ditampilkan 10 hasil iterasi yang juga mencakup iterasi dengan nilai GCV terkecil,

**Tabel 4.5** Pemilihan Titik Knot Optimum dengan Kombinasi Knot

GCV	$x_1$	$x_2$	$x_3$	$x_4$
35,56	21,364	79,811	97,067	70,313
37,29	21,364	79,811	97,066	70,316
			98,619	79,815
	21,364	46,547	97,066	38,651
17,45		76,484	98,619	67,149
		89,789		79,815
	21,364	79,811	97,066	70,316
39,37	26,749		98,619	79,815
	21,364	46,547	91,891	38,651
16,59	26,749	76,484	96,549	67,149
		89,789	98,619	79,815
	3,411	79,811	97,067	70,316
20,44	19,568	89,789		79,815
	26,749			
	3,411	79,811	91,891	70,313
18,83	19,568	89,789	96,549	
	26,749		98,619	
	<b>3,411</b>	<b>46,547</b>	<b>97,066</b>	<b>38,651</b>
<b>10,42</b>	<b>19,568</b>	<b>76,484</b>	<b>98,619</b>	<b>67,149</b>
	<b>26,749</b>	<b>89,789</b>		<b>79,815</b>
13,83	3,411	46,547	91,891	70,316
	19,568	76,484	96,549	79,815
	26,749	89,789	98,619	
12,73	3,411	46,547	91,891	38,651
	19,568	76,484	96,549	67,149
	26,749	89,789	98,619	79,815

Berdasarkan tabel diatas didapatkan nilai GCV terkecil yaitu sebesar 10,42 dengan kombinasi knot (3,3,2,3). Nilai titik knot untuk masing-masing variabel adalah sebagai berikut.

$$\begin{array}{lll} K_1 = 3,411; & K_2 = 19,568; & K_3 = 26,749; \\ K_4 = 46,547; & K_5 = 76,484; & K_6 = 89,789; \\ K_7 = 97,066; & K_8 = 98,619; & \\ K_9 = 38,651; & K_{10} = 67,149; & K_{11} = 79,815; \end{array}$$

#### 4.5 Pemilihan Model Terbaik

Setelah mendapatkan nilai GCV dari masing-masing pemodelan dengan satu knot, dua knot, tiga knot, dan kombinasi knot adalah membandingkan masing-masing nilai tersebut untuk mendapatkan pemodelan yang optimal/ terbaik. Berikut tabel nilai GCV dari masing-masing pemodelan dengan knot.

**Tabel 4.6** Nilai GCV Masing-Masing Knot

GCV	Jumlah Knot	Jumlah Parameter
35.56	1	9
13,66	2	13
12,73	3	17
10,42	Kombinasi Knot (3,3,2,3)	16 *

Berdasarkan model dengan satu, dua, tiga maupun kombinasi knot, model terbaik didapatkan dari nilai GCV paling minimum yaitu pada kombinasi knot(3,3,2,3) dengan nilai GCV sebesar 10,42 dan jumlah parameter sebanyak 16 parameter termasuk parameter  $\beta_0$ .

#### 4.6 Penaksiran Parameter Model Presentase Kasus Diare Pada Balita di Kota Surabaya

Penaksir parameter dari model terbaik yang didapatkan setelah melihat dari nilai GCV terkecil yaitu model kombinasi knot (3,3,2,3) adalah sebagai berikut;

$$\begin{aligned}\hat{y} = & \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+ + \hat{\beta}_{13}(x_1 - K_2)_+ + \hat{\beta}_{14}(x_1 - K_3)_+ + \hat{\beta}_{21}x_2 + \\ & \hat{\beta}_{22}(x_2 - K_4)_+ + \hat{\beta}_{23}(x_2 - K_5)_+ + \hat{\beta}_{24}(x_2 - K_6)_+ + \hat{\beta}_{31}x_3 + \\ & \hat{\beta}_{32}(x_3 - K_7)_+ + \hat{\beta}_{33}(x_3 - K_8)_+ + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_9)_+ + \\ & \hat{\beta}_{43}(x_4 - K_{10})_+ + \hat{\beta}_{44}(x_4 - K_{11})_+\end{aligned}$$

## 4.7 Signifikansi Parameter

Setelah mendapatkan model terbaik dengan nilai GCV terkecil maka dilanjutkan melakukan pengujian signifikansi parameter untuk mengetahui variabel prediktor apa saja yang berpengaruh signifikan terhadap model.

### 4.7.1 Uji Serentak

Pada uji serentak akan diketahui apakah variabel-variabel prediktor yang digunakan berpengaruh secara serentak terhadap model. Berikut hasil ANOVA untuk model regresi nonparametrik spline truncated.

**Tabel 4.7** ANOVA Model Regresi Spline Truncated

Sumber variasi	df	Sum of Square (SS)	Mean Square (MS)	$F_{hitung}$
Regresi	15	769,430	51,29	10,17
Error	15	75,640	5,04	
Total	30	845,07		

Berdasarkan hasil ANOVA seperti Tabel 4.7 diatas dapat diketahui nilai  $F_{hitung} (10,17) > F_{(0.05,15,15)} (2,40)$ . Maka dapat disimpulkan bahwa tolak  $H_0$  yang berarti minimal ada satu parameter yang signifikan terhadap presentase kasus diare balita di Kota Surabaya. Selanjutnya adalah melakukan pengujian secara individu untuk melihat parameter mana yang berpengaruh

signifikan terhadap presentase kasus diare balita di Kota Surabaya.

#### 4.7.2 Uji Individu

Untuk melakukan pengujian signifikansi parameter secara individu/ parsial dilakukan uji parsial. Berikut hasil pengujian parameter secara individu/parsial.

**Tabel 4.8** Hasil Pengujian Parameter Secara Individu

Variabel	Parameter	Estimator	T	P-value	Keputusan
$x_1$	$\beta_{11}$	4,1311	1,9085	0,0757	Tidak Signifikan
	$\beta_{12}$	-4,0084	-1,7965	0,0926	Tidak Signifikan
	$\beta_{13}$	-2,2518	-3,2742	0,0051	Signifikan
	$\beta_{14}$	4,2223	2,8774	0,0115	Signifikan
$x_2$	$\beta_{21}$	3,4561	2,2368	0,0409	Signifikan
	$\beta_{22}$	-4,0108	-2,5418	0,0226	Signifikan
	$\beta_{23}$	1,8023	7,7757	0,0000	Signifikan
	$\beta_{24}$	-4,3763	-8,4003	0,0000	Signifikan
$x_3$	$\beta_{31}$	1,5844	2,7903	0,0137	Signifikan
	$\beta_{32}$	-6,4780	-4,5469	0,0004	Signifikan
	$\beta_{33}$	5,9462	3,1267	0,0069	Signifikan
$x_4$	$\beta_{41}$	-2,1819	-1,6142	0,1273	Tidak Signifikan
	$\beta_{42}$	2,5765	1,8422	0,0853	Tidak Signifikan
	$\beta_{43}$	0,0303	0,1473	0,8849	Tidak Signifikan
	$\beta_{44}$	-2,0271	-3,4922	0,0033	Signifikan

Berdasarkan Tabel 4.8 terlihat bahwa terdapat beberapa parameter yang tidak signifikan terhadap presentase kasus diare pada balita di Kota Surabaya tahun 2015, karena ada nilai  $P_{\text{value}}$  pada beberapa parameter yang bernilai lebih dari  $\alpha(0,05)$  . Namun secara umum keempat variabel berpengaruh signifikan terhadap model karena masih ada parameter yang signifikan pada masing-masing variabel tersebut.

## 4.8 Pengujian Asumsi Residual

Model yang terbentuk dari analisis regresi nonparametrik nonparametrik Spline Truncated harus memenuhi beberapa asumsi antara lain residual harus identik, independen dan berdistribusi normal.

### 4.8.1 Asumsi Identik

Pengujian asumsi identik dilakukan untuk mengetahui apakah terdapat homogenitas dari varians residual. Untuk melakukan pengujian asumsi identik dapat menggunakan uji Glejser. Berikut merupakan hasil tabel ANOVA untuk uji Glejser.

**Tabel 4.9** Hasil Pengujian *Glejser*

<b>Sumber variasi</b>	<b>df</b>	<b><i>Sum of Square</i> (SS)</b>	<b><i>Mean Square</i> (MS)</b>	<b><i>F<sub>hitung</sub></i></b>
Regresi	15	17,09	1,13	0,903
Error	15	18,91	1,26	
Total	30	36,00		

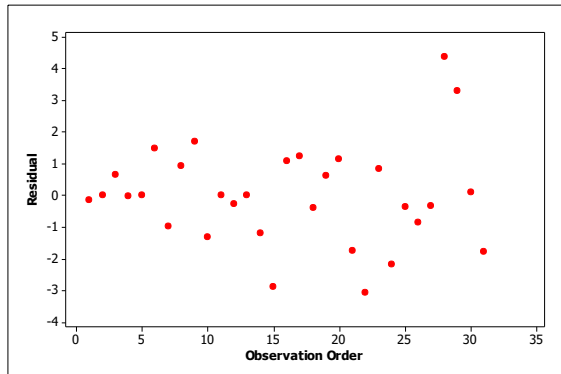
$$P_{\text{value}} = 0,576$$

Berdasarkan Tabel 4.9 didapatkan nilai  $F_{\text{hitung}} = 0,903 < F_{(0,05;15;15)} (2,40)$  serta nilai  $P_{\text{value}}$  yang bernilai  $0,576 > \alpha (0,05)$  maka dapat disimpulkan gagal tolak  $H_0$  yang artinya terjadi homogenitas pada varians residual sehingga asumsi identik terpenuhi.

### 4.8.2 Asumsi Independen

Pengujian asumsi independen dapat dilakukan dengan melihat nilai  $Z$  dan  $P_{\text{value}}$  pada *Run Test*. Berdasarkan output program didapatkan nilai  $Z (0,00) < Z_{\alpha/2} (1,96)$  dan nilai  $P_{\text{value}}$  sebesar  $0,812 > \alpha(0,05)$  sehingga gagal Tolak  $H_0$ . Hal ini menunjukkan bahwa asumsi independen terpenuhi. Selain itu asumsi independen dapat dibuktikan dengan pemeriksaan plot

antara residual dengan *observation order*. Berikut plot yang dihasilkan,

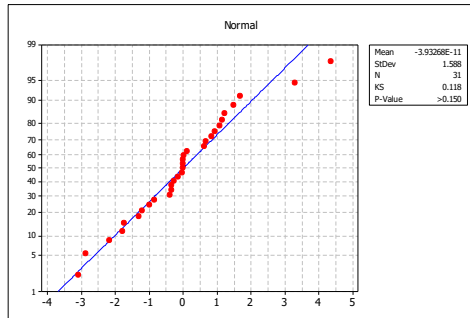


**Gambar 4.6** Scatterplot Antara Residual dengan *observation order*

Berdasarkan *scatterplot* antara residual dengan *observation order* diatas dapat diketahui persebaran residual terlihat acak. Hal ini menandakan tidaknya autokorelasi antar residual yang artinya asumsi independen terpenuhi.

#### 4.8.3 Asumsi Distribusi Normal

Pengujian asumsi residual berdistribusi normal dilakukan untuk mengetahui apakah residual sudah mengikuti distribusi normal atau tidak. Asumsi residual berdistribusi normal diperlukan karena dalam regresi nonparametrik spline truncated, selain varians residual harus homogen (asumsi identik) dan tidak ada autokorelasi antar residual (asumsi independen), residual juga harus berdistribusi normal. Salah satu cara yang dapat digunakan adalah menggunakan uji Kolmogorov Smirnov. Dengan perhitungan manual uji Kolmogorov Smirnov didapatkan nilai  $D = 0,118 < D_{tabel} (0,238)$  sehingga gagal tolak  $H_0$  yang berarti asumsi distribusi normal terpenuhi. Selain itu juga dapat melalui pemeriksaan dengan melihat *Normal Probability Plot Residual* seperti gambar berikut,



**Gambar 4.7** Plot Residual

Berdasarkan Gambar 4.7 terlihat bahwa nilai  $P_{\text{value}} > 0,150$  yang berarti lebih dari  $\alpha$  (0,05) maka dapat disimpulkan gagal tolak  $H_0$  sehingga residual dalam model tersebut sudah mengikuti distribusi normal.

#### 4.9 Intepretasi Model

Model terbaik dari regresi nonparametrik spline truncated pada kasus penyakit diare pada balita di kota Surabaya yaitu dengan menggunakan kombinasi knot (3,3,2,3). Nilai koefisien determinasi dari model tersebut adalah 91,04 persen. Hal ini dapat menjelaskan bahwa variabel presentase jumlah kasus diare pada balita di Kota Surabaya dapat dijelaskan oleh keempat variabel prediktor sebesar 91,04 persen. Sedangkan sisanya yakni 8,96 persen dijelaskan oleh variabel lain yang tidak terdapat dalam model. Berikut adalah model terbaik yang terpilih setelah memasukkan masing-masing nilai parameter dan knot.

$$\begin{aligned} \hat{y} = & -228,92 + 4,131x_1 - 4,008(x_1 - 3,411)_+ - 2,252(x_1 - 19,568)_+ + \\ & 4,222(x_1 - 26,749)_+ + 3,456x_2 - 4,011(x_2 - 46,547)_+ + \\ & 1,802(x_2 - 76,484)_+ - 4,376(x_2 - 89,789)_+ + 1,584x_3 - \\ & 6,478(x_3 - 97,006)_+ + 5,946(x_3 - 98,619)_+ - 2,182x_4 + \\ & 2,576(x_4 - 38,651)_+ + 0,03(x_4 - 67,149)_+ - 2,027(x_4 - 79,815)_+ \end{aligned}$$



Model tersebut dapat diinterpretasi sebagai berikut

1. Hubungan antara kepadatan penduduk ( $x_1$ ) dengan presentase kasus diare balita di Kota Surabaya ( $y$ ) dengan asumsi variabel lain dianggap tetap atau konstan adalah sebagai berikut;

$$\hat{y} = 4,131x_1 - 4,008(x_1 - 3,411)_+ - 2,252(x_1 - 19,568)_+ + 4,222(x_1 - 26,749)_+$$

$$= \begin{cases} 4,131x_1 & ; & x_1 < 3,411 \\ 0,123x_1 + 13,671 & ; & 3,411 \leq x_1 < 19,568 \\ -2,129x_1 + 57,738 & ; & 19,568 \leq x_1 < 26,749 \\ 2,093x_1 - 55,195 & ; & x_1 \geq 26,749 \end{cases}$$

Berdasarkan model tersebut dapat disimpulkan bahwa pada kecamatan dengan kepadatan penduduk kurang dari 3,411 ribu jiwa/km<sup>2</sup> dan apabila kepadatan penduduk naik 1 satuan maka presentase kasus diare balita akan naik 4,131 %. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Asemworo, Benowo, Pakal, Lakarsantri, Sambikerep.

Pada kecamatan dengan kepadatan penduduk antara 3,411 ribu jiwa/km<sup>2</sup> hingga 19,568 ribu jiwa/km<sup>2</sup> dan apabila kepadatan penduduk naik 1 satuan maka presentase kasus diare balita akan naik 0,123 %. Kecamatan yang termasuk dalam segmen ini adalah Sukomanunggal, Tandes, Genteng, Pabean Cantikan, Semampir, Krembangan, Bulak, Kenjeran, Gubeng, Rungkut, Tenggilis, Gunung Anyar, Sukolilo, Mulyorejo, Wonokromo, Karangpilang, Dukuh Pakis, Wiyung, Gayungan, Wonocolo, dan Jambangan.

Pada kecamatan dengan kepadatan penduduk antara 19,568 ribu jiwa/km<sup>2</sup> hingga 26,749 ribu jiwa/km<sup>2</sup> dan apabila kepadatan penduduk naik 1 satuan maka presentase kasus diare balita akan

turun 2,129 %. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Tegalsari, Bubutan, Tambak Sari, dan Sawahan.

Pada kecamatan dengan kepadatan penduduk lebih dari 26,749 ribu jiwa/km<sup>2</sup> dan apabila kepadatan penduduk naik 1 satuan maka presentase kasus diare balita akan naik 2,093 %. Kecamatan yang termasuk dalam segmen ini adalah Simokerto.

2. Hubungan antara presentase rumah tangga ber-PHBS ( $x_2$ ) dengan presentase kasus diare balita di Kota Surabaya ( $y$ ) dengan asumsi variabel lain dianggap tetap atau konstan adalah sebagai berikut;

$$\begin{aligned} \hat{y} &= 3,456x_2 - 4,011(x_2 - 46,547)_+ + 1,802(x_2 - 76,484)_+ - \\ &\quad 4,376(x_2 - 89,789)_+ \\ &= \begin{cases} 3,456x_2 & ; & x_2 < 46,547 \\ -0,555x_2 + 186,7 & ; & 46,547 \leq x_2 < 76,484 \\ 1,247x_2 + 48,875 & ; & 76,484 \leq x_2 < 89,789 \\ -3,129x_2 + 441,792 & ; & x_2 \geq 89,789 \end{cases} \end{aligned}$$

Berdasarkan model tersebut dapat disimpulkan bahwa pada kecamatan dengan presentase rumah tangga ber-PHBS kurang dari 46,547% dan apabila presentase rumah tangga ber-PHBS naik 1 satuan maka presentase kasus diare balita akan naik 3,46%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Semampir.

Pada kecamatan dengan presentase rumah tangga ber-PHBS antara 46,547% hingga 76,484% dan apabila rumah tangga ber-PHBS naik 1 satuan maka presentase kasus diare balita akan turun 0,555%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Sukomanunggal, Tandes, Asemworo, Pakal, Lakarsantri, Sambikerep, Tegalsari, Simokerto, Pabean Cantikan,

Krembangan, Kenjeran, Rungkut, Mulyorejo, Sawahan, Wonokromo, Karangpilang, Wonocolo, dan Jambangan.

Pada kecamatan dengan presentase rumah tangga ber-PHBS antara 76,484% hingga 89,789% dan apabila rumah tangga ber-PHBS naik 1 satuan maka presentase kasus diare balita akan naik 1,247%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Benowo, Genteng, Bubutan, Bulak, Tambaksari, Sukolilo, Dukuh Pakis, Wiyung, Gayungan.

Pada kecamatan dengan presentase rumah tangga ber-PHBS lebih dari 97,551% dan apabila rumah tangga ber-PHBS naik 1 satuan maka presentase kasus diare balita akan turun 3,129%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Gubeng, Tenggilis dan Gunung Anyar.

3. Hubungan antara presentase penduduk dengan akses jamban sehat ( $x_3$ ) dengan presentase kasus diare balita di Kota Surabaya ( $y$ ) dengan asumsi variabel lain dianggap tetap atau konstan adalah sebagai berikut;

$$\hat{y} = 1,584x_3 - 6,478(x_3 - 97,006)_+ + 5,946(x_3 - 98,619)_+$$

$$= \begin{cases} -1,584x_3 & ; & x_3 < 97,006 \\ -4,894x_3 + 628,404 & ; & 97,006 \leq x_3 < 98,619 \\ 1,052x_3 + 42,016 & ; & x_3 \geq 98,619 \end{cases}$$

Berdasarkan model tersebut dapat disimpulkan bahwa pada kecamatan dengan presentase penduduk dengan akses jamban sehat kurang dari 97,006% dan apabila penduduk dengan akses jamban sehat naik 1 satuan maka presentase kasus diare balita akan turun 1,584%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Asemworo, Bubutan, Simokerto, Semampir, Krembangan, Tenggilis, Sawahan.

Pada kecamatan dengan presentase penduduk dengan akses jamban sehat antara 97,006% hingga 98,619% dan apabila penduduk dengan akses jamban sehat naik 1 satuan maka presentase kasus diare balita akan turun 4,894%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Sukomanunggal, Benowo, Genteng, Tegalsari, Pabean Cantikan, Gubeng, Rungkut, Mulyorejo dan Gayungan.

Pada kecamatan dengan presentase penduduk dengan akses jamban sehat lebih dari 98,619% dan apabila penduduk dengan akses jamban sehat naik 1 satuan maka presentase kasus diare balita akan naik 1,052%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Tandes, Pakal, Lakarsantri, Sambikerep, Bulak, Kenjeran, Tambaksari, Gunung Anyar, Sukolilo, Wonokromo, Karangpilang, Dukuh Pakis, Wiyung, Wonocolo, dan Jambangan.

4. Hubungan antara presentase bayi dengan ASI eksklusif 0-6 bulan ( $x_4$ ) dengan presentase kasus diare balita di Kota Surabaya ( $y$ ) dengan asumsi variabel lain dianggap tetap atau konstan adalah sebagai berikut;

$$\begin{aligned} \hat{y} &= -2,182x_4 + 2,576(x_4 - 38,651)_+ + 0,03(x_4 - 67,149)_+ - \\ &\quad 2,027(x_4 - 79,815)_+ \\ &= \begin{cases} -2,182x_4 & ; & x_4 < 38,651 \\ 0,394x_4 - 99,564 & ; & 38,651 \leq x_4 < 67,149 \\ 0,424x_4 - 101,576 & ; & 67,149 \leq x_4 < 79,815 \\ -1,603x_4 + 60,206 & ; & x_4 \geq 79,815 \end{cases} \end{aligned}$$

Berdasarkan model tersebut dapat disimpulkan bahwa pada kecamatan dengan presentase bayi dengan ASI eksklusif 0-6 bulan kurang dari 38,651% dan apabila bayi dengan ASI eksklusif 0-6 bulan naik 1 satuan maka presentase kasus diare

balita akan turun 2,182%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Tandes.

Pada kecamatan dengan presentase bayi dengan ASI eksklusif 0-6 bulan antara 38,651% hingga 67,149% dan apabila bayi dengan ASI eksklusif 0-6 bulan naik 1 satuan maka presentase kasus diare balita akan naik 0,394%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Benowo, Lakarsantri, Sambikerep, Bubutan, Simokerto, Pabean Cantikan, Semampir, Krembangan, Bulak, Kenjeran, Tambak Sari, Gubeng, Rungkut, Mulyorejo, Sawahan, Wonokromo, Karangpilang, Dukuh Pakis dan Wiyung.

Pada kecamatan dengan presentase bayi dengan ASI eksklusif 0-6 bulan antara 67,149% hingga 79,815% dan apabila bayi dengan ASI eksklusif 0-6 bulan naik 1 satuan maka presentase kasus diare balita akan naik 0,424%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Asemrowo, Genteng, Tegalsari, Gunung Anyar, Gayungan, Wonocolo, dan Jambangan.

Pada kecamatan dengan presentase bayi dengan ASI eksklusif 0-6 bulan lebih dari 79,815% dan apabila bayi dengan ASI eksklusif 0-6 bulan naik 1 satuan maka presentase kasus diare balita akan turun 1,603%. Kecamatan yang termasuk dalam segmen ini adalah kecamatan Sukomanunggal, Pakal, Tenggilis daan Sukolilo.

Berikutnya akan disajikan skenario model optimis, *middle*, dan pesimis yang didapat dari model yang sudah optimum berdasarkan nilai GCV terkecil untuk memprediksi presentase diare balita di Kota Surabaya. Berdasarkan faktor-faktor yang memengaruhinya.

**Tabel 4.10** Prediksi Model Optimis, *Middle*, Pesimis

Presentase Diare Balita	$x_1$	$x_2$	$x_3$	$x_4$	Nilai Prediksi Presentase Diare Balita	Skenario Model
3,93	5,02	71,35	98,64	58,41	2,16	Optimis
2,31	24,20	82,45	98,75	67,99	2,60	
3,97	7,67	74,83	99,91	64,09	4,14	
4,20	11,13	77,43	99,38	67,97	5,82	<i>Middle</i>
7,33	11,73	55,35	98,19	40,23	6,75	
10,54	2,94	62,90	91,63	73,72	9,90	
13,11	11,49	65,35	99,82	70,68	12,48	Pesimis
16,02	7,22	80,24	97,81	75,56	15,24	
28,83	14,41	92,05	97,16	68,64	21,17	

Skenario model optimis menunjukkan kesuksesan suatu kecamatan di Kota Surabaya karena jumlah diare balita yang dijumpai relatif rendah. Model pesimis menunjukkan masih relatif tingginya diare balita yang dijumpai. Sedangkan skenario model *middle* berada pada *range* optimis dan pesimis. Untuk Model prediksi optimis dari presentase diare balita tersebut sebesar 2,16% dengan nilai presentase diare balita aktual yaitu sebesar 3,93% pada Kecamatan Rungkut. Model prediksi *middle* pada presentase diare balita tersebut sebesar 5,82% dengan nilai presentase diare balita aktual yaitu sebesar 4,20% pada Kecamatan Wonokromo. Model prediksi pesimis pada presentase diare balita tersebut sebesar 21,17% dengan nilai presentase diare balita aktual yaitu sebesar 28,83% pada Kecamatan Genteng.

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Kesimpulan yang dapat diperoleh dari penelitian tentang presentase kasus diare pada balita di Kota Surabaya adalah sebagai berikut.

1. Kecamatan yang memiliki presentase kasus diare pada balita tertinggi di Kota Surabaya tahun 2015 adalah kecamatan Genteng dengan presentase sebesar 28,83%. Sedangkan kecamatan dengan presentase kasus diare balita terendah terdapat pada kecamatan Tambaksari yaitu sebesar 2,31%. Untuk nilai varians dari presentase kasus diare balita di Kota Surabaya adalah sebesar 28,169. Hal ini menandakan nilai varians cukup besar sehingga persebaran kasus diare balita kurang merata di Kota Surabaya.
2. Model regresi nonparametrik Spline Truncated terbaik yang terpilih dalam kasus presentase diare balita di Kota Surabaya tahun 2015 adalah dengan kombinasi knot (3,3,2,3) dengan seluruh variabel prediktor signifikan terhadap variabel respon dengan nilai koefisien determinasi dari model tersebut adalah 91,04 persen. Berikut model yang digunakan;

$$\begin{aligned} \hat{y} = & -228,92 + 4,131x_1 - 4,008(x_1 - 3,411)_+ - 2,252(x_1 - 19,568)_+ + \\ & 4,222(x_1 - 26,749)_+ + 3,456x_2 - 4,011(x_2 - 46,547)_+ + \\ & 1,802(x_2 - 76,484)_+ - 4,376(x_2 - 89,789)_+ + 1,584x_3 - \\ & 6,478(x_3 - 97,006)_+ + 5,946(x_3 - 98,619)_+ - 2,182x_4 + \\ & 2,576(x_4 - 38,651)_+ + 0,03(x_4 - 67,149)_+ - 2,027(x_4 - 79,815)_+ \end{aligned}$$

3. Model prediksi optimis dari presentase diare balita tersebut sebesar 2,16% dengan nilai presentase diare balita aktual yaitu sebesar 3,93% pada Kecamatan Rungkut. Model prediksi *middle* pada presentase diare balita tersebut sebesar 5,82% dengan nilai presentase diare balita aktual yaitu sebesar 4,20% pada Kecamatan Wonokromo. Model prediksi pesimis pada presentase diare balita tersebut sebesar 21,17% dengan nilai presentase diare balita aktual yaitu sebesar 28,83% pada Kecamatan Genteng.

## 5.2 Saran

Saran yang dapat diberikan penulis pada penelitian kali ini adalah sebagai berikut,

1. Pemerintah Kota Surabaya harus memperhatikan kualitas kesehatan pada balita khususnya pada kasus diare dengan melakukan sosialisasi khususnya mengenai pentingnya berperilaku hidup sehat pada setiap rumah tangga dan pentingnya pemberian ASI eksklusif selama 0-6 bagi bayi karena berpengaruh signifikan terhadap presentase kasus diare di Kota Surabaya.
2. Penelitian selanjutnya dapat menambahkan kemungkinan titik knot pada setiap variabel nonparametrik karena pada penelitian ini hanya menggunakan 1, 2, 3 titik knot dan kombinasi knot.



## DAFTAR PUSTAKA

- Ayuningrum, F.V (2015). *Analisis Faktor Sanitasi Dan Sumber Minum Pada Insiden Diare Pada Balita Di Jawa Timur Dengan Regresi Logistik Biner*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Badan Pusat Statistik Kota Surabaya. (2015). *Statistik Kesehatan Kota Surabaya Tahun 2015*. Badan Pusat Statistik Kota Surabaya.
- Budiantara, I.N. (2000). Metode U, GLM, CV, dan GCV dalam Regresi Nonparametrik Spline. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6, 41-45.
- Budiantara, I.N. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.
- Depkes RI. (2007). *Riset Kesehatan Dasar (Riskesdas) Tahun 2007*. Badan Penelitian dan Pengembangan Kesehatan, Jakarta.
- Dinkes Jatim. (2015). *Profil Kesehatan Kota Surabaya Tahun 2015*. Dinas Kesehatan Kota Surabaya
- Draper, N.R. dan Smith, H. (1992). *Analisis Regresi Terapan*, Edisi Kedua, Alih Bahasa: Bambang Sumantri, PT Gramedia Pustaka Utama, Jakarta.
- Ernawati, Fitri (2015). *Pemodelan dan Pemetaan Kejadian Diare di Kota Surabaya Menggunakan Geographically Weighted Negative Binomial Regression*. . Surabaya: Institut Teknologi Sepuluh Nopember.
- Eubank, R. (1988). *Nonparametric Regression and Spline Smoothing*. Marcel Dekker Inc. New York.
- Gujarati, D. (2009). *Basic Econometrics (Ekonometrika Dasar)*. Alih bahasa: Sumarno Zain. Jakarta: Penerbit Erlangga.
- M.C Widjaja, (2002). *Mengatasi Diare dan Keracunan Pada Balita*. Jakarta: Penerbit Kawan Pustaka.
- Magdarina. (2011). *Morbiditas dan Mortalitas Diare pada Balita di Indonesia, Tahun 2000-2007*. Kementrian Kesehatan RI. Jakarta.

- Nisa', F.F. (2016). *Pemodelan Faktor-Faktor yang Memengaruhi Jumlah Kasus Tuberkolosis di Jawa Timur Menggunakan Regresi Nonparametrik Spline*. . Surabaya: Institut Teknologi Sepuluh Nopember.
- Nettina, Sandra M. (2001). *Pedoman Praktik Keperawatan. Penerbit Buku Kedokteran*. EGC, Jakarta.
- Pamungkas, W.S. (2013). *Linieritas Analisis Regresi*. Magister Manajemen Universitas Muhammadiyah Yogyakarta.
- Parashar DC (1993). *Methane emission estimate from Indian paddy fields*. Presented at the International Workshop Methane and Nitrous Oxide, Amersfoort. The Netherlands.
- Setiawan, & Kusini, D. E. (2010). *Ekonometrika*. C.V. Andi Offset.
- Wahba, G. (1990). *Spline Models For Observational Data*. Penerbit: Siam.
- Walpole, R.E. (1995). *Pengantar Metode Statistika*. Edisi Ketiga, Alih Bahasa: Bambang Sumantri. Jakarta: Penerbit PT Gramedia Pusaka Utama.
- WHO (2010). *World Health Statistics 2010: Causes of death*. France: WHO Press.

## LAMPIRAN

### Lampiran 1. Data Penelitian

Kecamatan	Y	$x_1$	$x_2$	$x_3$	$x_4$
Sukomanunggal	7,25	11,227	73,85	97,79	84,20
Tandes	4,96	9,590	63,40	99,61	36,54
Asemworo	10,54	2,849	61,31	91,55	72,76
Benowo	5,37	2,349	79,61	97,73	64,63
Pakal	3,32	2,215	52,67	99,23	88,26
Lakarsantri	14,22	3,364	57,19	99,97	62,72
Sambikerep	5,13	2,967	62,69	100	57,79
Genteng	28,83	11,865	89,72	97,08	67,74
Tegalsari	8,99	20,556	74,60	98,58	75,72
Bubutan	4,93	22,546	81,19	96,66	55,06
Simokerto	5,86	31,537	65,91	94,09	61,71
Pabean Cantikan	7,33	10,515	53,95	98,11	39,70
Semampir	12,69	17,807	44,33	92,30	65,59
Kremlangan	10,53	13,174	67,80	95,06	64,03
Bulak	3,87	6,609	83,66	99,00	53,11
Kenjeran	10,06	19,528	64,27	99,50	58,92
Tambaksari	2,31	23,466	80,36	98,66	67,10
Gubeng	7,09	16,518	92,83	98,45	55,43
Rungkut	3,93	5,917	69,54	98,55	57,65
Tenggilis	2,52	13,523	98,66	96,15	81,55
Gunung Anyar	3,45	6,589	94,96	99,59	69,05
Sukolilo	6,06	5,212	77,55	99,01	80,28
Mulyorejo	6,48	14,080	74,49	98,25	64,35
Sawahan	11,50	20,750	63,07	96,25	61,88
Wonokromo	4,20	9,657	75,47	99,29	67,08

**Lampiran 1. Data Penelitian (Lanjutan)**

Kecamatan	Y	$x_1$	$x_2$	$x_3$	$x_4$
Karangpilang	3,97	8,089	72,93	99,83	63,25
Dukuh Pakis	4,53	6,658	80,12	100	56,82
Wiyung	7,93	5,620	81,42	99,97	49,86
Gayungan	16,02	7,248	78,21	97,72	74,57
Wonocolo	13,11	12,196	63,69	99,73	69,76
Jambangan	6,34	11,409	76,35	99,73	75,05

**Lampiran 2. Program GCV 1 Knot**

```

GCV1=function(para)
{
data=read.table("d:/datadiare.txt")
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
dataA=data[, (para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk=
    length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot1=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot1[j,i]=a[j]
}
}
a1=length(knot1[,1])
knot1=knot1[2:(a1-1),]
aa=rep(1,p)
data1=matrix(ncol=m,nrow=p)

```

## Lampiran 2. Program GCV 1 Knot (Lanjutan)

```

data2=data[,2:q]
a2=nrow(knot1)
GCV=rep(NA,a2)
Rsqr=rep(NA,a2)
for (i in 1:a2)
{
  for (j in 1:m)
  {
    for (k in 1:p)
    {
      if (data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else
        data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
    }
  }
  mx=cbind(aa,data2,data1)
  mx=as.matrix(mx)
  C=pinv(t(mx)%%mx)
  B=C%%(t(mx)%%data[,1])
  yhat=mx% %B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsq[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx% %C% %t(mx)
  A1=(F-A)
  A2=(sum(diag(A1))/p)^2
  GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsq)

```

### Lampiran 2. Program GCV 1 Knot (Lanjutan)

```

cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqu dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (Rsqu)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsqu))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1
    knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="d:/output GCV1.csv")
write.csv(Rsqu, file="d:/output Rsqu1.csv")
write.csv(knot1, file="d:/output knot1.csv")
}

```

### Lampiran 3. Program GCV 2 Knot

```

GCV2=function(para)
{
data=read.table("d:/datadiare.txt")
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-1-para
dataA=data[, (para+2):q]
F=matrix(0, nrow=p, ncol=p)
diag(F)=1

```

### Lampiran 3. Program GCV 2 Knot (Lanjutan)

```

nk=
    length(seq(min(data[,para+2]),max(data[,para+2]
        ),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
a1=nrow(knot)
knot=knot[2:(a1-1),]
a2=nrow(knot)
z=(a2*(a2-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
for ( j in 1:(a2-1))
{
for (k in (j+1):a2)
{
xx=cbind(knot[j,i],knot[k,i])
knot1=rbind(knot1,xx)
}
}
knot2=cbind(knot2,knot1)
}
knot2=knot2[2:(z+1),2:(2*m+1)]
aa=rep(1,p)
data2=matrix(ncol=(2*m),nrow=p)
data1=data[, (para+2):q]
data3=data[,2:q]
a3=length(knot2[,1])
GCV=rep(NA,a3)

```

### Lampiran 3. Program GCV 2 Knot (Lanjutan)

```

Rsq=rep(NA,a3)
for (i in 1:a3)
{
for (j in 1:(2*m))
{
if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
for (k in 1:p)
{
if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else
data2[k,j]=data1[k,b]-knot2[i,j]
}
}
mx=cbind(aa,data3,data2)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C*%(t(mx)%*%data[,1])
yhat=mx*%*%B
SSE=0
SSR=0
for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx*%*%C*%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====\n")

```



### Lampiran 3. Program GCV 2 Knot (Lanjutan)

```

    cat("Nilai Knot dengan Spline linear 2
    knot", "\n")
cat("=====
=====","\n")
print (knot2)
cat("=====
=====","\n")
cat("Rsqr dengan Spline linear 2 knot","\n")
cat("=====
=====","\n")
print (Rsqr)
cat("=====
=====","\n")
cat("HASIL GCV dengan Spline linear 2 knot","\n")
cat("=====
=====","\n")
print (GCV)
s1=min(GCV)
cat("=====
=====","\n")
cat("HASIL GCV terkecil dengan Spline linear 2
    knot","\n")
cat("=====
=====","\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="d:/output GCV2.csv")
write.csv(Rsqr,file="d:/output Rsqr2.csv")
write.csv(knot2,file="d:/output knot2.csv")
}

```

### Lampiran 4. Program GCV 3 Knot

```

GCV3=function(para)
{
data=read.table("d:/datadiare.txt")

```

**Lampiran 4. Program GCV 3 Knot (Lanjutan)**

```

data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
F=matrix(0,nrow=p,ncol=p)
dataA=data[, (para+2):q]
diag(F)=1
nk=
length(seq(min(data[,2]),max(data[,2]),length.out=50)
)
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
knot=knot[2:(nk-1),]
a2=nrow(knot)
z=(a2*(a2-1)*(a2-2)/6)
knot1=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot2=rbind(rep(NA,3))
for ( j in 1:(a2-2))
{
for (k in (j+1):(a2-1))
{
for (g in (k+1):a2)
{
xx=cbind(knot[j,i],knot[k,i],knot[g,i])
knot2=rbind(knot2,xx)
}
}
}
}
}

```

#### Lampiran 4. Program GCV 3 Knot (Lanjutan)

```

knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[, (para+2):q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else
      data1[k,j]=data2[k,b]-knot1[i,j]
    }
  }
  mx=cbind(aa,data[,2:q],data1)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)

```

#### Lampiran 4. Program GCV 3 Knot (Lanjutan)

```

A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsqr)
r=max(Rsqr)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3
knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="d:/output GCV3.csv")
write.csv(Rsqr, file="d:/output Rsqr3.csv")
write.csv(knot1, file="d:/output knot3.csv")
}

```

#### Lampiran 5. Program GCV Kombinasi Knot

```

GCVkom=function(para)
{
  data=read.table("d:/datadiare.txt")
  data=as.matrix(data)

```

### Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)

```

p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("d:/x1.txt")
x2=read.table("d:/x2.txt")
x3=read.table("d:/x3.txt")
x4=read.table("d:/x4.txt")
x5=read.table("d:/x5.txt")
n2=nrow(x1)
a=matrix(nrow=5,ncol=3^5)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (s in 1:3)
{
m=m+1
a[,m]=c(i,j,k,l,s)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^5)
for (i in 1:3^5)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{

```

**Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)**

```

    if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]
  }
}
else
  if (a[i,1]==2)
  {
    gab=as.matrix(x1[,2:3])
    gen=as.matrix(cbind(data[,v],data[,v]))
    aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
    for (j in 1:2)
    for (w in 1:nrow(data))
    {
      if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]
    }
  }
  else
  {
    gab=as.matrix(x1[,4:6])
    gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
    aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
    for (j in 1:3)
    for (w in 1:nrow(data))
    {
      if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]
    }
  }
  if (a[i,2]==1)
  {
    gab=as.matrix(x2[,1] )
    gen=as.matrix(data[, (v+1)])
    bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
    for (j in 1:1)
    for (w in 1:nrow(data))
    {

```

### Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)

```

    if (gen[w,j]<gab[h,j]) bb[w,j]=0 else
bb[w,j]=gen[w,j]-gab[h,j]
  }
}
else
  if (a[i,2]==2)
  {
    gab=as.matrix(x2[,2:3] )
    gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)]))
    bb=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
    for (j in 1:2)
    for (w in 1:nrow(data))
    {
      if (gen[w,j]<gab[h,j]) bb[w,j]=0 else
bb[w,j]=gen[w,j]-gab[h,j]
    }
  }
  else
  {
    gab=as.matrix(x2[,4:6])
    gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)],data[, (
v+1)]))
    bb=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
    for (j in 1:3)
    for (w in 1:nrow(data))
    {
      if (gen[w,j]<gab[h,j]) bb[w,j]=0 else
bb[w,j]=gen[w,j]-gab[h,j]
    }
  }
  if (a[i,3]==1)
  {
    gab=as.matrix(x3[,1] )
    gen=as.matrix(data[, (v+2)])
    cc=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
    for (j in 1:1)
    for (w in 1:nrow(data))

```

**Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)**

```

{
  if (gen[w,j]<gab[h,j]) cc[w,j]=0 else
cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{
gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[, (v+2)],data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
  if (gen[w,j]<gab[h,j]) cc[w,j]=0 else
cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[, (v+2)],data[, (v+2)],data[, (
v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
  if (gen[w,j]<gab[h,j]) cc[w,j]=0 else
cc[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)

```



### Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)

```

for (w in 1:nrow(data))
{
  if (gen[w,j]<gab[h,j]) dd[w,j]=0 else
dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,4]==2)
{
  gab=as.matrix(x4[,2:3] )
  gen=as.matrix(cbind(data[, (v+3)],data[, (v+3)]))
  dd=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
  for (j in 1:2)
  for (w in 1:nrow(data))
  {
    if (gen[w,j]<gab[h,j]) dd[w,j]=0 else
dd[w,j]=gen[w,j]-gab[h,j]
  }
}
else
{
  gab=as.matrix(x4[,4:6])
  gen=as.matrix(cbind(data[, (v+3)],data[, (v+3)],data[, (
v+3)]))
  dd=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
  for (j in 1:3)
  for (w in 1:nrow(data))
  {
    if (gen[w,j]<gab[h,j]) dd[w,j]=0 else
dd[w,j]=gen[w,j]-gab[h,j]
  }
}
if (a[i,5]==1)
{
  gab=as.matrix(x5[,1] )
  gen=as.matrix(data[, (v+4)])
  ee=matrix(nrow=nrow(x1)*nrow(data),ncol=1)

```

**Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)**

```

for (j in 1:1)
for (w in 1:nrow(data))
{
  if (gen[w,j]<gab[h,j]) ee[w,j]=0 else
ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,5]==2)
{
  gab=as.matrix(x5[,2:3] )
  gen=as.matrix(cbind(data[, (v+4)],data[, (v+4)]))
  ee=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
  for (j in 1:2)
  for (w in 1:nrow(data))
  {
    if (gen[w,j]<gab[h,j]) ee[w,j]=0 else
ee[w,j]=gen[w,j]-gab[h,j]
  }
}
else
{
  gab=as.matrix(x5[,4:6])
  gen=as.matrix(cbind(data[, (v+4)],data[, (v+4)],data[, (
v+4)]))
  ee=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
  for (j in 1:3)
  for (w in 1:nrow(data))
  {
    if (gen[w,j]<gab[h,j]) ee[w,j]=0 else
ee[w,j]=gen[w,j]-gab[h,j]
  }
}
ma=as.matrix(cbind(aa,bb,cc,dd,ee))
mx=cbind(rep(1,nrow(data)),data[,2:q1],na.omit(ma))
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)

```

### Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)

```

B=C%*(t(mx)%*data[,1])
yhat=mx%*B
SSE=0
SSR=0
for (r in 1:nrow(data))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsqr=(SSR/(SSE+SSR))*100
MSE=SSE/p1
A=mx%*C%*t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}

if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
if (a[i,5]==1) splines=x5[,1] else
if (a[i,5]==2) splines=x5[,2:3] else
splines=x5[,4:6]
kkk=cbind(sp,spl,splin,spline,splines)
cat("=====", "\n")
print(i)

```

### Lampiran 5. Program GCV Kombinasi Knot (Lanjutan)

```
print(kkk)
print(Rsq)
}
write.csv(GCV,file="d:/output GCV kombinasi.csv")
write.csv(Rsq,file="d:/output Rsq kombinasi.csv")
}
```

### Lampiran 6. Program Penaksiran dan Signifikansi Parameter

```
parameter=function(alpha,para)
{
  data=read.table("d:/datadiare.txt")
  knot=read.table("d:/knotkom.txt")
  data=as.matrix(data)
  knot=as.matrix(knot)
  ybar=mean(data[,1])
  m=para+2
  p=nrow(data)
  q=ncol(data)
  dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3])
  dataA=as.matrix(dataA)
  satu=rep(1,p)
  n1=ncol(knot)
  data.knot=matrix(ncol=n1,nrow=p)
  for (i in 1:n1)
  {
    for(j in 1:p)
    {
      if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else
      data.knot[j,i]=dataA[j,i]-knot[1,i]
    }
  }
  mx=cbind(satu, data[,2],data.knot[,1:3],data[,3],
  data.knot[,4:6], data[,4], data.knot[,7:8], data[,5],
  data.knot[,9:11])
}
```

### Lampiran 6. Program Penaksiran dan Signifikansi Parameter (Lanjutan)

```

mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====", "\n")
cat("Estimasi Parameter", "\n")
cat("=====", "\n")
print (B)
n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/(SSR+SSE))*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
  cat("-----", "\n")
  cat("Kesimpulan hasil uji serentak", "\n")
  cat("-----", "\n")
  cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan", "\n")
  cat("", "\n")
}
else
{
  cat("-----", "\n")
  cat("Kesimpulan hasil uji serentak", "\n")
  cat("-----", "\n")
  cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan", "\n")
  cat("", "\n")
}

```

**Lampiran 6. Program Penaksiran Signifikansi Parameter (Lanjutan)**

```

}
#uji t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx))))
cat("-----","\n")
cat("Kesimpulan hasil uji individu","\n")
cat("-----","\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
  thit[i]=B[i,1]/SE[i]
  pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
  if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor
signifikan dengan pvalue",pval[i],"\\n") else
cat("Gagal tolak Ho yakni prediktor tidak signifikan
dengan pvalue",pval[i],"\\n")
}
thit=as.matrix(thit)
cat("=====","\n")
cat("nilai t hitung","\n")
cat("=====","\n")
print (thit)
cat("Analysis of Variance","\n")
cat("=====","\n")
cat("Sumber          df          SS          MS
Fhit","\n")
cat("Regresi          ",(n1-1)," ",SSR,"
",MSR," ",Fhit,"\\n")
cat("Error            ",p-n1," ",SSE," ",MSE,"\\n")
cat("Total            ",p-1," ",SST,"\\n")
cat("=====","\n")
cat("s=",sqrt(MSE),"          Rsq=",Rsq,"\\n")
cat("pvalue(F)=",pvalue,"\\n")

```

### Lampiran 6. Program Penaksiran dan Signifikansi Parameter (Lanjutan)

```
write.csv(res,file="d:/output uji residual.csv")
write.csv(pval,file="d:/output uji pvalue.csv")
write.csv(mx,file="d:/output uji mx.csv")
write.csv(yhat,file="d:/output uji yhat.csv")
write.csv(B,file="d:/output uji B.csv")
write.csv(thit,file="d:/output thit.csv")
}
```

### Lampiran 7. Program Uji Glejser

```
glejser=function(alpha,para)
{
data=read.table("d:/datadiare.txt")
knot=read.table("d:/knotkom.txt")
res=read.table("d:/resi.txt")
data=as.matrix(data)
knot=as.matrix(knot)
res=abs(res)
res=as.matrix(res)
rbar=mean(res)
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else
data.knot[j,i]=dataA[j,i]-knot[1,i]
}

```

### Lampiran 7. Program Uji Glejser (Lanjutan)

```

}
}
mx=cbind(satu, data[,2],data.knot[,1:3],data[,3],
data.knot[,4:6], data[,4], data.knot[,7:8], data[,5],
data.knot[,9:11])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
residual=res-yhat
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/SST)*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan atau terjadi heteroskedastisitas","\n")
cat("","\n")
}
else
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak
berpengaruh signifikan atau tidak terjadi
heteroskedastisitas","\n")
}

```



### Lampiran 7. Program Uji Glejser (Lanjutan)

```

cat("", "\n")
}
cat("Analysis of Variance", "\n")
cat("=====", "\n")
cat("Sumber      df      SS      MS
Fhit", "\n")
cat("Regresi      ", (n1-1), " ", SSR, "
", MSR, "", Fhit, "\n")
cat("Error      ", p-n1, " ", SSE, "", MSE, "\n")
cat("Total      ", p-1, " ", SST, "\n")
cat("=====", "\n")
cat("s=", sqrt(MSE), "      Rsq=", Rsq, "\n")
cat("pvalue(F)=", pvalue, "\n")
}

```

### Lampiran 8. Output Penaksiran dan Signifikansi Parameter

```

=====
Estimasi Parameter
=====
      [,1]
[1,] -228.92071513
[2,]  4.13108387
[3,] -4.00841154
[4,] -2.25179075
[5,]  4.22225533
[6,]  3.45609881
[7,] -4.01080240
[8,]  1.80227159
[9,] -4.37630449
[10,] 1.58439512
[11,] -6.47801893
[12,]  5.94622892
[13,] -2.18194424
[14,]  2.57646912
[15,]  0.03032996
[16,] -2.02705602

```

## Lampiran 8. Output Penaksiran dan Signifikansi Parameter (Lanjutan)

-----  
Kesimpulan hasil uji serentak  
-----

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

-----  
Kesimpulan hasil uji individu  
-----

Tolak Ho yakni prediktor signifikan dengan pvalue 0.01429577

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue  
0.07565286

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue  
0.09257761

Tolak Ho yakni prediktor signifikan dengan pvalue 0.00512288

Tolak Ho yakni prediktor signifikan dengan pvalue 0.01150907

Tolak Ho yakni prediktor signifikan dengan pvalue 0.04090776

Tolak Ho yakni prediktor signifikan dengan pvalue 0.02256012

Tolak Ho yakni prediktor signifikan dengan pvalue 1.220096e-06

Tolak Ho yakni prediktor signifikan dengan pvalue 4.695512e-07

Tolak Ho yakni prediktor signifikan dengan pvalue 0.01372462

Tolak Ho yakni prediktor signifikan dengan pvalue 0.0003855294

Tolak Ho yakni prediktor signifikan dengan pvalue 0.006928654

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue  
0.127323

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue  
0.08530384

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue  
0.8848765

Tolak Ho yakni prediktor signifikan dengan pvalue 0.00327561

=====  
nilai t hitung

=====  
[,1]

[1,] -2.7701061

[2,] 1.9084869

### Lampiran 8. Output Penaksiran dan Signifikansi Parameter (Lanjutan)

```
[3,] -1.7965038
[4,] -3.2741949
[5,] 2.8774405
[6,] 2.2368476
[7,] -2.5418342
[8,] 7.7756813
[9,] -8.4003074
[10,] 2.7903375
[11,] -4.5469471
[12,] 3.1267039
[13,] -1.6141804
[14,] 1.8421871
[15,] 0.1472747
[16,] -3.4921897

Analysis of Variance
=====
Sumber      df      SS      MS      Fhit
Regresi     15  769.4314  51.29542  10.17218
Error       15   75.64078  5.042719
Total       30  845.0722
=====
s= 2.2456 Rsq= 91.04919
pvalue(F)= 2.683009e-05
```

### Lampiran 9. Output Uji Glejser

-----
Kesimpulan hasil uji serentak
-----
Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan atau tidak terjadi heteroskedastisitas
Analysis of Variance

```
=====
Sumber    df    SS    MS    Fhit
Regresi   15  17.09788  1.139858  0.9037175
Error     15  18.91949  1.2613
Total     30  36.01737
=====
s= 1.123076 Rsq= 47.4712
pvalue(F)= 0.5764301
```

## Lampiran 10. Surat Pernyataan Data



### PEMERINTAH KOTA SURABAYA DINAS KESEHATAN

Jalan Jemursari No. 197 Surabaya 60243  
Telp. (031) 8439473, 8439372, 8473729 Fax. (031) 8483393

Surabaya, 9 Mei 2017

Nomor : 072 / 1759 / 436.7.2 / 2017  
Sifat : Biasa  
Lampiran :  
Hal : Selesai Pengambilan Data

Kepada  
Yth. Ketua Jurusan Statistika  
Institut Teknologi Sepuluh  
Nopember

di-

#### SURABAYA

Memperhatikan surat Ketua Jurusan Statistika Institut Teknologi Sepuluh Nopember nomor : 003669/IT2.VI.1.3/TU.00.09/2017 tanggal : 19 Januari 2017 perihal pada pokok surat tersebut diatas, kami informasikan bahwa mahasiswa FMIPA-ITS atas nama Muhammad Syauqi Khudzaifi telah selesai pengambilan data di Dinas Kesehatan Kota Surabaya.

Demikian atas perhatian dan kerjasamanya disampaikan terima kasih.

a.n. KEPALA DINAS  
Sekretaris,



Nanik Sukristina, SKM  
Pembina

NIP. 197001171994032008

## BIODATA PENULIS



Penulis dengan nama lengkap Muhammad Syauqi Khudzaifi dan akrab dipanggil Dafi lahir dari pasangan Ahmad Khariri dan Khofsotun di Kota Jepara pada tanggal 29 Oktober 1994 dan merupakan anak kedua dari dua bersaudara. Penulis yang lahir dan besar di Kota Jepara ini telah menempuh pendidikan formal di SD Negeri Panggang 1 Jepara (2001-2007), SMP Negeri 1 Jepara (2007-2010), SMA Negeri 1 Pati (2010-2013), dan menyelesaikan pendidikan di Departemen Statistika Institut Teknologi Sepuluh Nopember Surabaya. Selama kuliah penulis aktif dalam kegiatan organisasi kampus antara lain menjadi Staf Departemen Dalam Negeri HIMASTA-ITS periode 2014-2015 dan Kepala Biro Minat dan Bakat HIMASTA-ITS periode 2015-2016 serta menjadi Steering Committee Bina Cinta Statistika HIMASTA-ITS 2015. Selain itu penulis juga aktif dalam perlombaan karya tulis mahasiswa seperti menjadi Ketua pada Program Kreativitas Mahasiswa pada bidang pengaduan sosial dan pernah didanai oleh Dikti sebanyak dua kali. Dalam menjalani kehidupan penulis mempunyai prinsip “Berusahalah menjadi pewarna dalam hal kebaikan pada kehidupan orang lain” karena sebaik-baik manusia adalah orang yang bermanfaat bagi orang lain. Apabila ada suatu pertanyaan atau ingin berdiskusi tentang Tugas Akhir, dapat menghubungi melalui telepon 085 812 090 525 atau melalui email: [muhammadsyauqik@gmail.com](mailto:muhammadsyauqik@gmail.com)